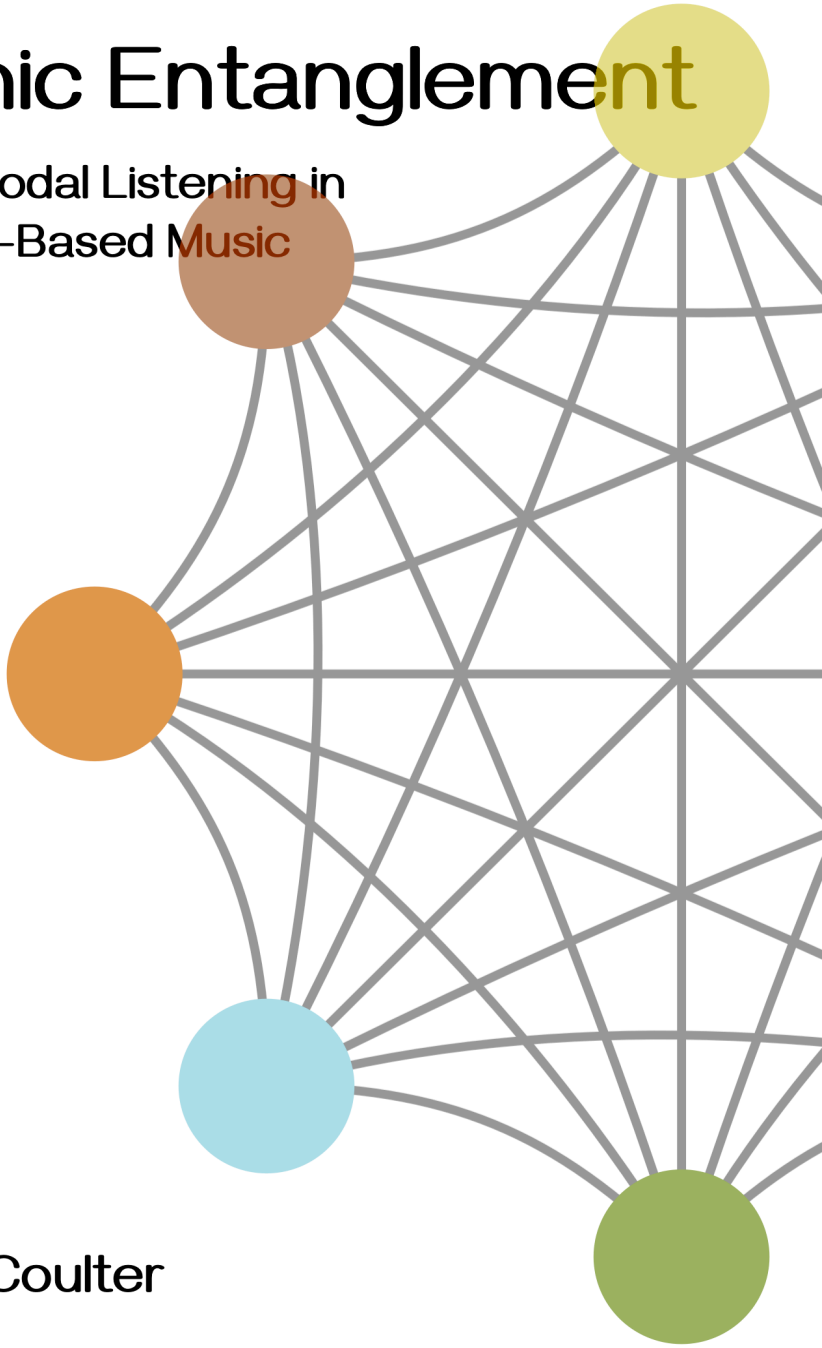
Sonic Entanglement

Multimodal Listening in Sound-Based Music



John Coulter

Sonic Entanglement

Multimodal Listening in Sound-Based Music

John Coulter

Name of publisher: ETTE Publications
Address of publisher: 32 Swiftsure Street Collingwood Tasman 7073 New Zealand
Telephone. +64 (0)21 181 7302

Author website: www.eartotheearth.nz

Copyright © 2026 by John Coulter
ISBN 978-1-0672920-0-3 (paperback)
ISBN 978-1-0672920-1-0 (hardback)
ISBN 978-1-0672920-2-7 (eBook)

Cover design and figure 2.4 by Charlie Budd
Figure 6.4 reprinted with permission from the author
All other figures reconstructed or authored by John Coulter

Acknowledgements

I would like to thank my partner Sheridan Gray for the personal support she has offered me, for her contributions to the book through conversations, the developing of ideas, direct writing, and the editing skills she has brought to the project. I would also like to thank my good friends and fellow sonic artists Chris Cree Brown, John Cousins, Gerardo Dirié and David Rylands for the sounding-board and expert feedback they provided me with over the five years it took to write. Thanks also to my international colleagues Simon Emmerson, Gary Kendall, Barry Truax, and John Young for the advice and support they offered me in preparing the final draft. Special thanks to Charlie Budd for designing the final figure and book cover.

Contents

Preface

This book offers composers and sonic artists a powerful set of integrated concepts for discussing and creating sound-based music. The innovative approach departs from previous stepwise, sequential ways of conceptualising the listening process and encourages creative individuals to think multimodally in their artmaking practices. Based on a collection of multisensory modes of listening, the model serves to simplify and unite numerous extant theories from the domains of electroacoustic music studies, acoustic ecology, psychology, and neuroscience. Eight interwoven modes of listening (modes of auditory attention) are identified, representing the: expressive, formal, referential, visual, gestural, environmental, spatial, and structural qualities of sound. New insights are offered through the concept of entanglement – the way in which all modes are simultaneously activated during the ordinary listening process, managed by different areas within the brain, and developed through crosstalk between layers of mental activity.

The title of the book, Sonic Entanglement describes the dynamic interplay of concurrent listening modes that emerge through the creation and reception of sound-based music. The term 'sound-based music' is used to define 'the art form in which the sound, that is, not the musical note, is its basic unit' (Landy 2007: 17). Applied to completed works from a range of genres and categories (including acousmatic music, visual music, performance-based works with live electronics, and interactive installation), the model provides higher levels of scrutiny and depth of discussion through its unified theoretical framework and expansive vocabulary. The principle is also relevant to the auditory experience of everyday life, as well a range of other creative disciplines (including fine arts, film, dance, theatre, and architecture). While several of the illustrative examples are drawn from these wider artforms, the majority are focussed on genres and categories of sonic art. Therefore, the book is recommended for advanced students, scholars, and established sonic artists who are interested in understanding and discussing the functionality of their own, and other's sound-based works of art.

Several important references in the book are sourced from the period 1990-2010. This was a significant time in history as it is the era in which the language of sound-based musical composition was developing at pace. Work of that time provides a scaffold on which our understanding of the listening experience rests. Of particular interest is the work of Pierre Schaeffer (1966), Hildegard Westerkamp (1974), Murray Schafer (1977, 1994), Barry Truax (1984, 1994, 2008, 2012), Denis Smalley (1986, 1996, 1997, 2007), Simon Emmerson (1986, 2007), Katherine Norman (1996), and more recently Gary Kendall (2010a, 2010b, 2014, 2016) and John Young (2016). Their theories and models are revered by the international sonic arts community and although their suppositions have been added to, they have not been supplanted. The conceptual framework, outlined in this book aims to combine these numerous foundational and supporting texts into a single working model.

Chapter One: Modes of Listening gives an overview of how listening has been theorised within this field previously. The chapter presents and expands on the key texts that define the modes of listening within the context of experiencing music and sonic art. Initially the work of Schaeffer (1966), Smalley (1986, 1996, 1997), and Emmerson (1986, 2007) set the foundations for this discussion. Smalley (1996) proposed three main relationship types between the sounds and the composer/listener: *reflexive, interactive,* and *indicative* – essentially relating to the expressive, formal, and referential qualities of sound. He adds that aesthetic attitudes, as well as experience/inexperience in listening to art music can influence individual listeners' experiences through shifts in modal *focus*. Westerkamp (1974), Schafer (1977), Truax (1993, 2012), Norman (1996), et al. propose the addition of *environmental,* and *visual,* (or quasi-visual), and *spatial* dimensions to the listening experience. Detail is then added to the referential (*what*) qualities of sound through Smalley's (1996) nine indicative fields: gesture, utterance, behaviour, energy, motion, object/substance, environment, vision, and space.

Chapter Two: Multimodal Listening distils modes of listening research into one working model incorporating eight interwoven fields: expressive, formal, referential, visual, gestural, environmental, spatial, and structural. The chapter follows Kendall's (2010a, 2010b, 2014, 2016)

lead in considering all modes of listening to be in simultaneously operation. This offers new perspectives on the research presented in chapter one with endorsements from Miranda (2010), Bregman (1993), Reiss Jones and Yee (1993), Snyder (2000), et al. The intersection of four contributing fields of research electroacoustic music studies, acoustic ecology, neuroscience, and psychology presents further clarity and insight into the complex way in which humans experience audio-based works of art. Of key significance is a re-definition of the listening experience as one that is *multimodal* and *multisensory.* While Smalley's (1996) *reflexive*, *interactive*, and *indicative* relationships are retained (as *expressive, formal,* and *referential* modes of listening) equal weight is assigned to the *visual* (or *quasi-visual*), *gestural* (body-centred), *environmental, spatial,* and *structural* aspects of sound.

Chapters 3-10 expand on the eight interwoven modes of listening presented in chapter two with a view to converting the observations to artmaking principles and strategies. While none of the modes function in isolation, this deductive approach allows deep investigation into associated subtopics. The resulting insights highlight the *entanglement* of all modes of listening. We find that the listening experience (together with the practice of sonic artmaking) relies on multimodal and multisensory *interdependence*.

Chapter Three: Musical Expression explores how emotions and feelings are commonly experienced through music and sound. The chapter is concerned with the way listeners *feel about* and make meaning of what they hear from an emotional perspective. It also addresses how musicians tend to think and act *during* the music-making process and how they might consider their own bodies and musicianship skills as *conduits* for expression. The chapter initially offers an expansion of Smalley's (1996) *reflexive* relationship through considering a range of emotions that might be commonly experienced through music and sound. It is observed – when listening - that basic types of emotion are experienced early, before complex emotional meanings develop. The composer or performer's tendency to intentionally *imbue* their sounds with various forms of emotion and meaning is also discussed, together with the perception of real and imaginary *sources* of such expressions.

Chapter Four: Formal Musical Qualities draws an overt relationship between traditional formal musical qualities (such as melody, harmony and rhythm) and Snyder's (2000) five image schemas - which may be described as mental patterns derived from bodily and perceptual experiences that help in making sense of the world. Beginning with Smalley's (1996) *interactive relationship*, it addresses the formal qualities of sound, focussing on the *pitch-time lattice* as a means of highlighting specific sonic characteristics that are found in both musical materials and in sounds drawn from the environment. Snyder (2000) specifies five *image schemas*: 1) up and down, 2) centrality, 3) motion-linkage-causation, 4) linearity: pathways and goals, and 5) containment: inside and outside. Image schemas, while relevant to all modes of listening, may be seen to highlight specific patterns displayed by abstract musical materials.

Chapter Five: Referential Meaning in Music explores the nature of sound-based meaning and expands on methods of managing referential sounds in sonic artworks. Initially the chapter focusses on Smalley's (1996) *indicative* relationship, exploring the nine indicative fields - highlighting the sonic characteristics that suggest association with sound-producing sources. Several contemporary theories concerning external semiotics (Molino, Underwood and Ayrey, 1990; Windsor, 1995; Atkinson, 2007; and Kim, 2008) are examined to advance understanding of referential *meaning* and its constituent parts within the context of experiencing music and sonic art. The active role of metaphor in the assignment of meaning is also investigated with a focus on pairing with Synder's five image schemas (up and down, centrality, motion-linkage-causation, linearity: pathways and goals, and containment: inside and outside). Finally, conceptual blending is examined as a means of exploring how metaphorical meaning might emerge within musical contexts. The second part of the chapter presents three strategies that provide methods of effectively managing referential meaning within any given sonic artwork: *augmentation of meaning-based qualities, justification of technique* and *conceptual blending as meaning*.

Chapter Six: Audiovisual Relationships explores how visual (or quasi-visual) associations, both real and imagined, play an important role in the understanding of sounds. The chapter presents a combination model

based on the findings of Chion (1994) Coulter (2010), Basanta (2017) and Garro (2005, 2012) - models concerning the artificial synthesis of audiovisual relationships in mixed-media artmaking contexts. Both *isomorphic* (corresponding in form and relations) and *concomitant* (naturally accompanying or associated) relationships are examined, first within the narrow context of on-screen sounds and images (for example, film or visual music), then in a wider context (for example in live performance settings), and the role these audiovisual relationships play in the multimodal appreciation of sonic artworks.

Chapter Seven: Gestural Associations highlights the important role human *gesture* has to play in the multimodal appreciation of sounds. As a starting point, Emmerson's (2007) local-field relationship is presented and elaborated on within the context of *mixed works* (incorporating live instrumental performance with live electronics). We examine the four associated spaces: event, stage, arena, and landscape, resulting in the emergence of several advanced artmaking strategies (relevant to the production of music and sonic art in performance-based settings). The united model explains how music is 'indelibly stamped with its bodily origins' (Clarke and Davidson 1998), *spreading out* from event space (where human gesture is evident), passing through the domains of stage and arena and further still to elicit associations with landscape and the environment.

Chapter Eight: Environmental Connections concerns active engagement with and study of the natural environment. The chapter evidences how sonic artists (particularly) are not confined to the desk or the studio, but tend to be active in the field, highlighting this direct relationship. From a wider perspective, it investigates an aesthetic appreciation of sound championed by the World Soundscape Project (WSP) constituting important concepts such as: 'taking the time to listen' (Westerkamp 2019), 'soundscapes' and 'soundmarks' (Schafer 1994), and 'soundscape composition' (Truax 2008), thus extending the musical paradigm into the realm of the environmental and awe-inspiring - while at the same time exposing the fragility and uniqueness of natural objects, events, and silences within environmental space. From a narrow perspective, several specific audio-based creative practices are examined, including stereo

and multichannel field recording, spectral transformation (the use of software) in the field, and environmentally activated instruments (both acoustic and sensor-based).

Chapter Nine: Spatial Forms proposes the construction of a theoretical model based on a set of archetypal spatial forms that contribute to the organisation of sound and the structuring of artworks. The chapter extends on Smalley's (2007) treatise in which he redefines and deconstructs the listening experience as a series of interacting *spatial forms*. Several other concepts and models relating to space are examined including Hall's four social zones (1966), Bachelard's intimacy and immensity (1964), Emmerson's local/field relationship (2007), Snyder's image schemas (2000), and Munt's heterotopias (2002). Common to all is the way in which human beings tend to appreciate space as stored memories. Space helps us to explain emotions, abstract musical qualities, sound objects and events, visual images, proprioceptive sensations, our location within the environment, and structural forms. The practice of multichannel acousmatic music is then examined, providing a narrow context for the exploration of these principles in a sub-practice of sonic arts that has come to be known as *spatial composition.*

Chapter Ten: Structural Concepts is concerned with the *organisation of sound.* This is relevant to both the *development of component structures* and *structural integration* phases in the creation of sonic artworks. These principles include moment form, morphic form, narrative, continuity, image schemas, the referential-abstract continuum, auditory Gestalten, wholeness, conceptual blending, and peak musical experiences. The wide focus of the chapter examines how these concepts contribute to the consolidative phenomenon of *conceptual blending,* unifying component structures to achieve structural integration. These strategies are proposed as a highly generalised range of artmaking techniques relevant to the management of structure in all musical contexts. A narrow finding of the chapter is the conclusion that *peak musical experiences* 'cannot be commanded' (Maslow 1962, also quoted by Green 2022: 45); however, composers and sonic artists may engineer situations in which *happy accidents* might occur – ideally leading to the emergence of these elusive qualities.

The final section *A Reduction of Fundamental Principles* offers composers and sonic artists a succinct reference guide for applying the key principles outlined in each chapter to analyses of their own creative works. The summary also specifies the theoretical maximums of two, three, four, five, six, seven, and eight-way relationships between the eight modes of listening identified - providing both the producers and listeners of sound-based music with an unprecedented range of simultaneous perspectives on the functionality of sonic materials.

Chapter One: Modes of Listening

How do we listen to artful sound? This chapter addresses modes of listening that are activated within the narrow context of experiencing music and sonic art - and the wider context of everyday life. The discussion will be limited to the *listening experience: in response to information received through sound-only sources*. (Although it will not be limited to *hearing* alone, as other senses also play a role in the appreciation of sound). Listening will initially be investigated through *the acousmatic* (sound-only) lens. An explanation of this term is first offered:

> A Pythagorean term reintroduced in 1955 by Jérôme Peignot which considers the "distance which separates sounds from their origin", i.e., an audio-only presentation of sound common to electroacoustic music. For some, the term is very precise and refers specifically to this listening situation. However, the term has gained wider usage, in describing a genre, which, to a large extent derives from the Musique Concrète tradition and is founded upon this listening situation. (EARs website accessed 12:12:24)

In applying this description, we might consider *the acousmatic* set of conditions as an ideal setting to investigate artful expression and communication through sound (alone). Acousmatic music is also well suited to the discussion of *modes of listening*, as the same psychological processes tend to be active in our everyday lives and in the experience of making sense of sonic art, as Gary Kendall explains:

> But before focusing on electroacoustic listening, we need to consider the broader question of what listeners do in everyday life when making sense of the world around them. This question is far more significant for electroacoustic music than for traditional acoustic music because the expansive acoustic palette of electroacoustic music incorporates the sounds of the everyday world – not just the sounds but, by immediate extension, the listening strategies, knowledge of auditory patterns and history of auditory associations that are inextricably linked to everyday listening. (Kendall 2010a: 63)

There are numerous contributors to the topic of *modes of listening*. These include methods of analysis proposed by the experts of electroacoustic

music and acoustic ecology, sound for film production techniques, and empirical studies carried out by researchers within the domain of cognitive psychology. As Barry Truax clarifies:

> With the traditional disciplines, what is most striking is the way in which the study of sonic phenomena has been fragmented across nearly all areas of academic discourse. Each area proceeds from its own theoretical models and methodology, using its own terms and language, essentially getting the "local picture" correct but ignoring the landscape (or soundscape) as a whole. (Truax 1993: 4)

It seems there has been much said about *how we listen*. A complete picture is, only now, beginning to emerge to create a 21st-century understanding of listening as a *multi-modal trans-sensory experience*. Katherine Norman iterates, 'Our ordinary listening is itself a complex, multi-layered activity of which hearing is but a part' (Norman 1996: 4). Active modes of listening can differ between individuals and are not easily defined as a finite set. However, we can specify a range of multi-sensory modes that are in common use. Some are concerned with:

- basic feelings of pleasure and displeasure
- the appreciation of music through the understanding of a code
- information gathering and causality
- aesthetic expectations or attitudes adopted by the listener
- associations with the environment
- associations with visual (or quasi-visual) images
- associations with bodily movements and feelings
- associations with spatial forms
- structural associations (making sense of multiple events)
- technological aspects of production

A broad model of *how we listen to artful sound* will emerge at the end of *Chapter Two: Multimodal Listening*. For now, the focus of discussion will concentrate on historical perspectives surrounding the topic. Gaining a full understanding of the common modes of listening and how they are triggered is essential for any composer or sonic artist who wishes to gain

greater access to the communicative potential of sound. There are several foundational theories and models that can support this understanding: Nattiez's tripartite model, the Schaefferian listening modes, Emmerson's Grid, and Smalley's tripartite model. By including these in the discussion as a starting point, we can come to a much deeper understanding of the complexity of human audition (the listening experience).

Nattiez's Tripartite Model

In 1990 Nattiez (and Molino) proposed the following tripartite analysis model as a detailed musical analysis method. In essence, it approaches musical appreciation from three standpoints:

1) the *poietic* – the producers' perspective.
2) the *esthesic* – the position of the receiver(s); and
3) the *trace* – the immanent structures of the work itself.

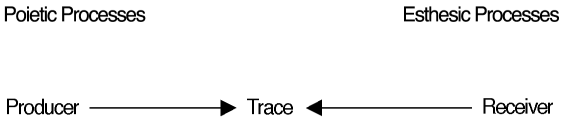


Fig. 1.1: Nattiez's tripartite analysis model (Nattiez 1990: 43).

The model can accommodate wide multi-sensory perspectives and variable production-reception processes (such as those found in mixed media art, film production, and performing arts); however, for the traditional musical experience, there are three main points of view: the author's, the performer's, and the listener's. As Nattiez states:

> Their relationship to one another is varied in the extreme, sometimes contradictory, sometimes confused …. In fact, composer, performer, and listener can without undue exaggeration be regarded not only as three types or degrees of relationship to music, but also as three successive states of specialisation. (Nattiez 1990: 31)

Individual perspectives of the composer, performer, and listener carry a specific set of concerns that can radically impact on the listening experience. Composers, performers, and listeners listen *for* different qualities within the music through focussing on *specific modes of listening*. This process of selective attention highlights and masks aspects of the sonic texture to such an extent that different listeners tend to *hear* different qualities within the same musical experience.

A feature of the model is the specification of both implicit and explicit methods of analysis under the categories of the *poietic* and the *esthesic*, alongside analysis of the neutral level (which concerns 'only the immanent configurations of the work'). This flexible approach allows the musical materials to be contemplated from multiple standpoints, resulting in the emergence of multiple types of analysis, understandings, and divergent messages. This results in five types of analyses, with a sixth condition that iterates the need to include *poietic*, *trace* and *esthesic* elements in any analysis. Details concerning these six methods are as follows (*ibid* 1990: 140-3).

I) **Immanent analysis [The Trace]:** This is a type of analysis that, working with an implicit or explicit methodology, tackles only the immanent configurations of the work (the measurable attributes of sound).

II) **Inductive poietics**: One can proceed from an analysis of the neutral level [the trace] to drawing conclusions about the poietic. For this particular instance I suggest the term inductive poietics …. We observe so many recurrent procedures within a work or body of works that we find it hard to believe that 'the composer had not thought of them'.

III) **External poietics**: In the third case the situation is reversed. The musicologist takes a poietic document – letters, plans, sketches – as his or her point of departure, and analyses the works in the light of this information.

IV) **Inductive esthesics**: constitutes the most common case, primarily because most analyses wish to style themselves up as the collective consciousness of listeners, and decree 'that is what one hears'. (An expert listener, often the composer, insisting on a specific type of reaction from their audience).

V) **External esthesics**: One can, on the other hand, begin with information collected from listeners, to attempt to understand how the work has been perceived.

VI) To these five analytical situations we need to add a sixth, one that appears simple, but in reality, is extremely complex. This is the case in which an immanent analysis is equally relevant to the poietic as to the esthesic… It necessitates pinning down the exact nature of the connection between the neutral and the poietic and the neutral and the esthesic.

A key feature of Nattiez's tripartite model is that it does not attempt to simplify the musical experience but instead considers the subjective needs and concerns of the listener. This gives rise to multiple understandings and interpretations that can commonly arise from the same musical materials.

The Schaefferian Listening Modes

Pierre Schaeffer, in delineating the listening process, identified four common modes that play roles in the process of attending music and sonic art (Schaeffer, 1966). Dozens of scholars from the field of music, (particularly electroacoustic music), have subsequently extended on and tested his treatise within the context of experiencing and composing sonic art. These experts include Denis Smalley, who is widely regarded as one of the leading authorities on the subject. For the purposes of this discussion, I have chosen to elaborate freely on Schaeffer's four modes with reference to Smalley (1996) who provides an excellent summary of his four modes of listening.

The first of Schaeffer's modes, named *Écouter*, concerns *information-gathering*. As humans we are preoccupied with the messages carried by/within sound for reasons related to navigation and safety. What is that sound? Where is it coming from? These are questions, from an evolutionary perspective, that must be answered quickly to survive in a hostile world. When sounds are experienced in mode one, they act *referentially*. The message-carrying attributes of the sound point to something else beyond the sound itself (for example, a living thing, a mechanical action, or a natural event). A mode one response automatically results in the assignment of meaning, which may also involve the decoding of sound (as is the case with spoken communication). Materials of this latter type often attract the parallel

term *semantic* (relating to meaning in language). The information gleaned from a listening experience remains highly subjective as it is intrinsically related to each individual listener's life experience (the languages they speak, their culture, and the sounds they have come to know through their lifetimes). As Smalley explains:

> We interpret these messages from the spectromorphology of the sound based on our accumulated knowledge of people, occurrences, environment, and the sounds associated with these experiences. In mode one, sounds are an index to a network of associations and experiences; we are concerned with causality; it is a question of living and acting in the world, ultimately of survival (Smalley 1996: 78).

From a compositional perspective, mode one (*Écouter*) may be used to tell stories or to create referential associations within musical contexts. Smalley also reminds us that *all sounds*, no matter how abstract, contain some form of referential meaning. No sound is exempt from a mode one interpretation.

The second mode, named *Ouïr* is concerned with *sounds that impinge on the listener's consciousness* - often fleeting moments in which loud, transient, or information-rich sounds dominate our attention. Sounds such as gunshots (exhibiting pronounced attention-grabbing qualities) are automatically highlighted, while all other sounds are often masked. From a compositional perspective, both referential and abstract sounds can operate in this way (shouts and screams, doors opening/closing, drum hits or impacts, bell sounds, etc.) These sounds may be *used* to emphasise important moments in any given musical artwork.

Lennox, Vaughan and Myatt (2001) present a set of sound types that might be considered attention-grabbing in the way intended by Schaeffer's mode two (*Ouïr*). They state that fast recognition processes have developed in humans in response to 'ambience labels' that are primarily contained in the upper frequencies of sound. These high frequencies also present the basic information required for object recognition to occur (mode one). They write:

> The 'what' representation yielded in the audio mode [Schaeffer's mode one] is different from, complementary to, and as richly detailed as that derived from the visual mode. Whilst form is not as well rendered, material composition, structure, type-of 'what' and size-of 'what' often are (Lennox, Vaughan and Myatt 2001: 4).

Lennox, Vaughan and Myatt's research provides an understanding of the way in which mode two and mode one might operate together. It seems that the attention-grabbing qualities of mode two may very well quicken the assignment of meaning that has been referred to as a mode one response. Lennox, Vaughan and Myatt outline a *recognition process model* based on a special attention-grabbing sonic quality they refer to as 'comingness' (a roll-on of high frequencies that are perceived as rushing towards the listener). Their research shows that this subset of sounds is the most attention-grabbing of all. This results in the following series of rapid recognition factors that occur in the early stages of the aural perception process: 1) quality of 'comingness', 2) what it is, and 3) where it is. The existence of this fast stage 1-2-3 process makes good sense from an evolutionary perspective, given that for countless eons, the need to flee from attacking predators was ever present. These sounds tend to be extremely memorable and, thus, also support the *structure* of sonic artworks.

Schaeffer's mode three *Entendre* encapsulates the *abstract qualities of sound*. It involves a passive or intentional appreciation of pitch-based and/or rhythm-based attributes of a sound for the sole purpose of aesthetic enjoyment. Sounds produced by birds, for example, are often appreciated for their melodic, harmonic, or rhythmic qualities. While mode one (*Écouter*) recognition factors - such as: bird, type of bird, male/female, etc. - may remain present, they are often suspended or relegated to the background to allow the sound's internal abstract characteristics to be appreciated. An important distinction between modes one and three is that the *referential qualities* of sound act as signs, pointing away from the sound itself (towards its meaning), whereas the *formal* qualities of sound may be conceptualised as intrinsic.

Attending to sounds in mode three (*Entendre*) can also result in the identification of more extended patterns, giving rise to a more sustained and complex musical experience. This concept is housed by Schaeffer's mode four *Comprendre*. Modes three and four can therefore be considered as early and later stages of an aesthetically focused continuum of listening experience. While modes one and two are not altogether absent, the primary concern is one of perusal and enjoyment. Again, the exact nature of what constitutes a musical pattern remains highly subjective as prior listening experience and musical education can radically impact on musical appreciation. As Smalley states:

> In mode four we find musical works, which are purposely created for listening. They can loosely be regarded as specially selected sounds encoded in a structure that has a sense and meaning for those listeners who share the code (though we may be very hard pressed to say what that sense or meaning might be) (Smalley 1996: 79).

Mode four (*Comprendre*) encompasses the *recognition of patterns* that operate both vertically and horizontally across the musical pitch-time lattice. These include forms and structures housed within individual works (for example: rising/falling of melodic patterns, harmonic progressions, textural expansions/contractions, rhythmic inversions, augmentations, and the known structures of popular and classical genres such as strophic, arch, and sonata form). From a compositional perspective, mode four: *Comprendre* is where the hand of the composer is most often revealed. The clever ways in which they draw the listener's attention to certain important patterns often becomes part of the enjoyment of the listening experience.

Table 1.1 summarises Schaeffer's four modes, referencing their primary characteristics and descriptions of associated listening experiences.

Mode	Characteristics	Listeners' experience
Mode one: Écouter	*Information-gathering*	Listeners preoccupied with identifying 'what' and 'where' the sound is.
Mode two: Ouïr	*Attention-grabbing*	Listeners' attention occupied by something striking. Urgent need for mode one to be activated.
Mode three: Entendre	*Abstract/formal qualities*	Listener experiences aesthetic enjoyment and perusal of sound
Mode four: Comprendre	*Recognition of extended musical patterns*	Listener experiences appreciation of the vertical/horizontal patterns of intentional musicality

Table 1.1: Schaeffer's Four Modes

Smalley iterates that none of the modes operate exclusively. He speaks of 'modal shading and shifts' within a listening experience that mixes and crosses all four modes depending on the qualities of the musical materials and the reactions of listeners. Composers may take advantage of this multimodal musical appreciation by deliberately steering listeners toward specific modes to emphasise musical qualities. Smalley specifies the 'essence' of the Schaefferian modes as contained in the mode one-three relationship. In an earlier publication (published in Emmerson, 1986), Smalley addressed this phenomenon in the following statement:

> All sounds possess this dual potential – the abstract and concrete aspects of sound – and all musical structures are balanced somewhere between the two, although exactly how they are found to be balanced can vary greatly among listeners. (Smalley 1986: 64)

The 'abstract aspects of sound' (also referred to as *formal musical qualities*) may be understood as mode three (*Entendre*), whereas 'the concrete aspects of sound' (also referred to as referential qualities) are bound to mode one (*Écouter*). Using Smalley's terms, the two qualities may be imagined as opposing ends of an abstract-concrete continuum

reflecting all listening experiences. Some listeners may experience a sound as primarily abstract (focussed on formal musical qualities), while others (listening to the same sound) may experience it as primarily concrete (focussed only on referential qualities). Most listeners; however, will tend to experience a 'balance' both formal and referential qualities simultaneously.

Emmerson's Grid

Simon Emmerson (1986) extends on Smalley's abstract-concrete continuum to include two methods of syntax (a term borrowed from language, meaning the arrangement of words or sounds to create well-formed sentences). He specifies how sonic artists/composers might manage the function and structure of sounds in sonic artworks and how these techniques might be interpreted by listeners. His use of terms will now be explained with reference to figure 1.2.

Abstracted syntax refers to methods of coaxing-out or identifying developmental pathways that are contained within the sounds themselves. For example, a vocal sample featuring the spoken words 'loop this phrase' might naturally and justifiably be open to developmental techniques that include multiple repetitions and other loop-based transformational techniques, whereas a vocal phrase of undefined content might not be as convincing if subjected to the same series of treatments.

Abstract syntax refers to developmental ideas that are *not* present or contained within the materials themselves (extra-musical ideas that are drawn from elsewhere). An example of *abstract syntax* might be the use of sonata form as a structuring device. Most composers and sonic artists tend to employ *abstract syntax* techniques at some stage within the compositional process, particularly in the latter stages. Observationally, formal musical sound categories tend to be more open to abstract syntax techniques than referential sound categories.

The terms *aural discourse* and *mimetic discourse* correspond directly to,

and are substitute terms for, Smalley's *abstract* and *concrete* qualities of sound as previously explained. *Aural discourse dominant* therefore refers to compositional approaches focussed on the *formal musical qualities* of sound, and *mimetic discourse dominant* refers to a focus on *referential* sounds. Emmerson considers sounds (particularly recorded sounds) *mimetic* due to their representative qualities.

An important feature of Emmerson's grid is that it clearly illustrates that *musical discourse* contains *both* aural and mimetic discourse (both formal and referential qualities). In his view, *all* sounds are part of the musical palette, in preference to delineating between musical and non-musical sounds. Surprisingly, this division of sound (between musical sounds and non-musical sounds or noise) is still common in some conservative musical sectors.

Abstract syntax	1	4	7
Combination of abstract and abstracted syntax	2	5	8
Abstracted syntax	3	6	9
	I: Aural discourse dominant	II: Combination of aural and mimetic discourse	III: Mimetic discourse dominant

MUSICAL DISCOURSE

Fig. 1. 2 Emmerson's Grid

The terms that eventually gained popular use by the international sonic arts community are drawn from both Smalley's (1986) and Emmerson's (1986) contributions. A summary is offered in table 1.2.

Popularised Term	*Parallel Terms*	*Meaning*
Abstract sounds	Aural discourse Formal Qualities Schaeffer's mode 3	The formal musical qualities of sound
Referential sounds	Concrete sounds Mimetic discourse Schaeffer's mode 1	The associative meaning of the sound
Abstracted syntax	Observed rules	Development strategies suggested by the sounds themselves
Abstract syntax	Applied rules	Development strategies drawn from elsewhere

Table 1.2: Summary of Terms

Smalley's Tripartite Model

in 1996 Smalley extended his thinking regarding the *'dual potential'* of sound, adding an additional mode of listening associated with basic emotional responses to music and sound. From this process, his tripartite model was born, encompassing 'indicative, reflexive and interactive relationships' (Smalley 1996: 82). The model provides sonic artists and composers with a more detailed overview of the message-bearing, sensational/emotional, and formal musical qualities of sound. It achieves this by delineating a set of sonic fields, defining the functional language-bearing qualities of music and sound. The model is well known throughout the international sonic arts community and remains one of the most detailed methods of identifying *sonic characteristics* within sound-based musical contexts.

Smalley used the term *autocentricity* (originally proposed by Ernest Schachtel) to encompass the *reflexive relationship* as a subject-centred sense focused on 'basic responses and feelings of pleasure and displeasure.' (Smalley, 1996: 81). Smalley explains that 'autocentric attitudes to music and sound are common. At its most primal, musical autocentricity is concerned with positive/negative emotional reaction to

sounds.' (ibid). The companion term *allocentric* is described as a perceptual mode that is 'object-centred in that it involves perceiving something independent of the perceiver's needs. The relationship between the qualities we perceive and any feelings of pleasure or displeasure we may have, is less direct or less pronounced. Indeed, pleasure or displeasure in the manner of autocentric reaction may even be irrelevant.' (ibid). Smalley goes on to state that 'allocentric attitudes to sounds encompass the apprehension of musical structure (Schaeffer's mode four), and our appreciation of sounds outside musical contexts in the manner of Schaeffer's mode three.' (ibid).

Smalley then proposed a synthesis of Schaeffer's and Schachtel's modes resulting in the formation of his tripartite model. I have chosen to present Smalley's submission in his own words as it offers the most succinct and exacting description of his model that is available. He writes:

> The term 'relationships' is preferred to 'modes', as it hints at impermanence and shifts in perception. The entire sound-field must be accounted for since any sound can become part of an electroacoustic work. The following three relationship-types are retained. *The indicative relationship* represents Schaeffer's mode one: sound as message, or information about environmental action, events, occurrences. It is object-centred and can be either actively or passively apprehended depending on whether the sound impinges on our consciousness or whether we are actively seeking it (waiting for a signal, for example). The *reflexive relationship* is taken from Schachtel's autocentric mode. It is subject-centred and is concerned with basic emotional responses to the object of perception. The object has no real identity separate from the subject's emotion. This type of relationship can be active or passive but has a strong tendency towards passivity since there is little or no exploration of the object of perception, only response to it. The *interactive relationship* embraces Schaeffer's modes three and four, and Schachtel's allocentric mode. It involves an active relationship on the part of the subject in continuously exploring the qualities and structure of the object. In this sense it can be regarded as interactive. No pre-formed expectation is sought by the listener. The interactive relationship embraces structural hearing, aesthetic attitudes towards music and sounds, and Schaeffer's concept of *reduced listening.* It is possible to have an interactive relationship with any sound, although naturally some sounds will be more fruitful than others. The

> indicative and reflexive relationships are the most prevalent. As far as listening to music is concerned the reflexive relationship popularly dominates, and as far as Western art music is concerned there should ideally be a partnership between the reflexive and interactive. Clearly an interactive relationship is necessary for the fullest appreciation of Western art music. All three relationships are subject to perceptual shifts which are within the focal gift of the subject, although not all listeners are equally gifted. (Smalley 1996: 82)

Most significantly, in the same article, Smalley expands on the indicative relationship, identifying nine contributing fields: gesture, utterance, behaviour, object/substance, energy, motion, environment, vision, and space. The first three fields, he specifies as 'archetypal' (presenting the idea of brain functions that may have developed in the early stages of human evolution). He notes that gesture and utterance provide the traditional models for music, with behaviour operating in ways analogous to certain types of 'human relationship, observed relationships among things or objects, or human-object relations.' (Smalley, 1996: 83).

The remaining six fields (object/substance, energy, motion, environment, vision, and space) specify other essential qualities and multisensory associations to present a detailed and highly descriptive combination model. In breaking down and specifying the contributing indicative fields, Smalley was able to identify ways to control the strength of referential association allowing *meaning* to be interpreted in far less rigid ways than was previously possible. This extensive paradigm will be examined at depth in *Chapter Five: Referential Meaning in Music*.

In the late 1990s, Smalley's tripartite model, together with rapidly developing music technology seemed to offer sonic artists unprecedented levels of theoretical and technical precision that might be applied in the creation of groundbreaking sonic artworks. Table 1.3 presents his model in succinct form:

Indicative Relationship (referential associations)	Reflexive Relationship (emotional responses)	Interactive Relationship (formal musical qualities)
Nine indicative fields: Gesture Utterance Behaviour Object/Substance Energy Motion Environment Vision Space	A range of common feelings evoked through music: e.g. amusement, joy, eroticism, beauty, relaxation, sadness, dreaminess, triumph, anxiety, scariness, annoyance, defiance, and feeling pumped up	Categories based on traditional music analysis: e.g. note, harmonics, timbre, intervallic structure, vertical harmony, dynamics, scales and modes, melody, harmonic progression, rhythm, key, duration, tempo, time signature, metre, pulse and pace, cadence, form, and structure

Table 1.3 Smalley's Tripartite Model

Smalley iterates that numerous modes of listening may be active at one time, that modes are not exclusive, and that attention can easily shift between modes during the listening process. Although he identifies just three main relationship types (as indicated), he also mentions that *other* (undefined) modes of auditory attention may play roles in the appreciation of music and sound. For example, *technological listening* – 'when a listener 'perceives' the technology or technique behind the music rather than the music itself' (Smalley 1997: 109). This is significant, as it reveals that Smalley does not consider the model to be comprehensive, rather, that individual listening experiences tend to evoke a wide range of listening modes as part of the ordinary process of human audition.

Other Important Contributions

Barry Truax (1984) recontextualised the listening experience as one that is inherently related to the soundscape of the natural world, suggesting that the binary division between information-carrying and abstract qualities of sound may be less pronounced in this context. He proposes

three levels of attention, the first being 'background listening' which can be extended further in the context of technology to 'distracted listening' (for example, accompaniment media, background music, etc.), the second, 'listening-in-readiness' (impacting attention), and the third, 'listening-in-search' (focussed listening) which includes a perusal and aesthetic enjoyment of naturally occurring sounds. The listening-in-search mode can also be extended through the employment of technology to become 'analytical listening', encompassing much of the electroacoustic music domain. (Truax 1984: 16-22). Truax's thesis on *soundscape listening* (primarily employing background, listening-in-readiness and listening-in-search modes) might be described in essence in the words of Katherine Norman as 'reflective'. She states:

> Reflective listening occurs when a listener does not simply identify a sound source but engages in a "perceptual interpretation" that balances memory and imagination. (Norman 1996: 250).

Expanding on this principle, in a more recent article, Truax suggests that sounds are interpreted though both the 'inner complexity' of sound organisation and the 'outer complexity' of relationships in the real world. He states:

> In other words, both inner and outer complexity inform our understanding of sound. Further, sound is not merely information exchange but is capable of creating relationships between listeners and their environment in a dynamic process of embodied cognition. This mediating role of sound creates the possibility of the 'acoustic community' (Truax 2012: 194)

Another important contribution concerns the conversion of the *sound object* put forward by Pierre Schaeffer to the *sound* (or sonic) *event*. The following extract is sourced from the Electroacoustic Resources Site (EARS).

> The sound event, like the sound object, is defined by the human ear as the smallest self-contained particle of a soundscape. It differs from the sound object in that the latter is an abstract acoustical object for study, while the sound event is a symbolic, semantic or structural object for study, and is

> therefore a non-abstractable point of reference, related to a whole of greater magnitude than itself. Source - R. Murray Schafer (1977)

Through this description, we can see that the sound event is one that bonds well with the subjective experience of living and acting in the world. Schafer also championed other key concepts relating to the reflective appreciation of sound in the environment. These include the *soundmark* – the sonic equivalent of landmark, and the *soundwalk* – 'an exploration of the *soundscape* of a given area using a score as a guide' (Schafer 1994: 213).

Hildegard Westerkamp articulated the concept of the *soundwalk* in her early essay 'Soundwalking'. The document was developed in 1974 as part of her work with the World Soundscape Project (WSP) of which Schafer and Truax were also a part. The definition and accompanying set of practical instructions is still in common use today (more than 50 years later) She states:

> A soundwalk is any excursion whose main purpose is listening to the environment. It is exposing our ears to every sound around us no matter where we are. … Unless we listen with attention, there is a danger that some of the more delicate and quiet sounds may pass unnoticed… (Westerkamp 1974: 1)

In applying the listening strategies articulated by Truax, Schafer, and Westerkamp to the current context, the role of the environment (one of Smalley's nine fields) is highlighted in the registration and understanding of sounds. The qualities carried by background environmental sounds alone, are enough, it seems, to activate several multisensory modes of listening simultaneously - including *quasi-visual* associations (often based on memories) that help to build important *spatial* relationships in the minds of listeners. This topic of listening in the environment will be expanded further in *Chapter Eight: Environmental Connections.*

Katherine Norman (1996) presents a model for *multi-layered listening* suggesting the dynamic employment of a range of listening strategies,

with individual subjectivity playing a major role in the assignment of meaning. She uses the terms 'referential' and 'reflective' to encompass the primary qualities of sound, affirming that both experiences are contextualised within an amassed field of individual experience and knowledge. Her approach highlights the radical differences in meaning that can be attained by different individuals when listening to electroacoustic music, presenting both uncertainty and opportunity within the artform. As Simon Emmerson states: 'she argues– almost in archetypal terms– that engaging with personal memory opens up greater possibilities of generating meanings for all listeners' (Emmerson 2015: 72).

Other ecological models of perception (Windsor 1995, 2000) and important additions from the field of semiotics (Kim 2010, et al) also do much to break down some of the divisions between Smalley's three main categories *reflexive, interactive,* and *indicative.* As these contributions relate mostly to the assignment of meaning, they will be discussed in further detail in *Chapter Five: Referential Meaning in Music.*

Aesthetic Attitudes

At this point in the discussion, it may be poignant to address aesthetic attitudes, as they also play a role in steering what individual listeners are *listening for* (steered by preconceived expectations). Within the context of Western art music, this type of aesthetic shading has been referred to as *schools of thought*. Venturing into a full analysis of the *schools* that harbour and promote these aesthetic attitudes within high art music-making contexts is beyond the remit of this book; however, it is important to note that these attitudes tend to be sustained by adherence to one or more of the relationships currently under discussion (reflexive, interactive, or indicative).

Historically, the French electroacoustic music aesthetic tends to be weighted towards the interactive relationship (Schaeffer's modes three and four) and, as such, employs the notion of *reduced listening* (training one's ears to ignore mode-one message-bearing information). The British

sonic arts aesthetic by comparison tends to promote a more balanced approach, making use of both referential and abstract materials, which brings with it its own set of opportunities and constraints. Such biases and shadings promote focussed attention on specific modes *of listening* often at the exclusion of others. Aesthetic attitudes highlight the *voluntary* (or socially steered) way in which any of the identified modes of listening might be seen to take precedence over others.

Experienced and Inexperienced Listeners

We must also take into consideration experience, and inexperience, in the practice of listening to artful sound. For *inexperienced* listeners, commonly adopted modes of listening tend to be culturally informed by other forms of music they are more familiar with and thus, music *outside* of the boundaries of their *normal listening experience* can be challenging. In scenarios where listeners are new to sound-based music (featuring indicative, reflexive and interactive qualities), the listening experience can produce a range of unintended strong referential associations. This presents a problem to the sonic artist, as it may transform their intentions and limits the communicative reach of their works to audiences who are more experienced in listening to electroacoustic music. Even then, it is likely that the aesthetic attitudes adopted by experienced listeners will steer their appreciation of the work. Various experts from the domain of electroacoustic music have presented approaches to resolve this problem. Leigh Landy has proposed what he calls the 'something to hold onto' factor. He asks:

> In such circumstances, what types of musical means are there to allow a new listener, particularly a less experienced one, to enter the world of a sound-based musical work? In other words, the goal here is the identification of useful musical devices that support the listening experience. (Landy 2007: 30)

Landy answers this question in several direct ways which include the use of narrative and/or dramatological devices, as well as gesture-to-sound relationships supplied within the context of live performance.

Concluding Comments

Gaining a detailed understanding of how human beings listen to and interpret sound is logically foundational to composers and sonic artists. Beginning with Schaeffer's (1966) initial four modes of listening, we have explored how the process of listening becomes relational. Smalley (1996) proposed three main relationship types: *reflexive, interactive,* and *indicative,* which, in essence, relate to the *emotional*, *formal*, and *referential* qualities of music and sound (respectively). It has also been put forward, that aesthetic attitudes, as well as experience/inexperience can influence individual listeners' experiences through shifts in modal *focus*. Other listening strategies proposed by Truax (1984, 2012), Schafer (1994), Westerkamp (1974) Norman (1996), et al. have added *environmental, visual,* and *spatial* associations to the list. Finally, Smalley's (1996) identification of nine indicative fields: gesture, utterance, behaviour, energy, motion, object/substance, environment, vision, and space, not only substantiates others' claims but, does much to position the *human body* at the centre of the listening experience. These historical perspectives on the topic of *modes of listening* combine to present a total listening experience as one that is essentially *multimodal* and *multisensory.* This sets the stage for further investigation of multimodal listening strategies, and reconciliation with contemporary research concerning the psychology of human audition. This will be the focus of *Chapter Two: Multimodal Listening.*

Chapter Two: Multimodal Listening

In recent years, new modes of listening have come to light that contain associations with other senses that further assist in understanding the message-bearing and abstract qualities of sound (beyond audio-only modes of listening). These modes are commonly referred to as multimodal or multisensory listening strategies. They embrace relationships with the environment, vision, proprioception (body-centred sensations), and spatial awareness. While Smalley (1996) has already identified the message-bearing sonic characteristics responsible for forming these relationships (the *nine indicative fields:* gesture, utterance, behaviour, object/substance, energy, motion, environment, vision and space), the focus of the current chapter is to reconsider each category as a multisensory mode of listening. This change in thinking has logical underpinnings – as Gary Kendall iterates:

> For the everyday mind, meaning is essentially multimodal, and it is important to remind ourselves that the meaning we make of auditory experience does not arise from within the context of auditory perception alone (Kendall 2010a: 68)

Where the previous chapter aims to explain *audio-only* modes of listening, this chapter combines established multisensory modes of listening (drawn from electroacoustic music studies and acoustic ecology), with research from the fields of psychology and neuroscience that identify both procedures and areas within the brain responsible for processing auditory events. As Miranda states:

> Essentially, the brain is a predictive organ, which strives to find structure in sensory information (or to impose it). In order to do this efficiently, it needs to make abstractions to fuel the relentless processes of making internal representations of the world. Behind these processes there is an impelling force to organise sensory information, which is driven by the physiological nature of our brain and its own evolving internal representations, or models, of the world. (Miranda 2010: 21)

The multisensory modes of listening that have already been identified in chapter one will now be briefly expounded.

Visual Relationships

A modern understanding of the listening experience invariably includes some form of visual registration – even if that appreciation is reduced to what is sometimes referred to as a *quasi-visual* experience (when an *imaginary image* is produced in the mind of the listener). Some sonic artists go to extreme measures to reduce visual stimuli of any kind - insisting on total blackout, blindfolds, and even concealing the location of loudspeakers - for the purpose of maximising such quasi-visual *imagination*. In settings involving the live performance of music, it is now known that the physical presence of a performer carries with it a powerful set of visual and social cues that can command artistic communicability. This can either support or detract from accompanying musical materials. When it comes to *music for film,* audiences expect to hear music and sound in its supporting role for moving images. Here, audio materials perform important functions, but it is a domain where visual materials dominate. Conversely, within the digital-media-based genre of *visual music* (where the artificial synthesis of audiovisual relationships underpins the domain), there is a preconception that certain types of visual materials (primarily abstract) will be *driven by* sounds, and that both audio and visual streams will play active roles in the delivery of artistic content. From this very brief survey it is apparent that the functionality of audiovisual relationships differs depending on the musical genre and category; however, it is also clear that the majority of listening experiences are in fact *audiovisual*, irrespective of whether sounds are intentionally accompanied by visual images. This important relationship, concerning both imaginary and real visual images is examined at depth in *Chapter Six: Audiovisual Relationships.*

Gestural Associations

Gestural associations operate within the wider category of referential associations (relationships concerning *the source or cause of sounds*). The subset is defined through the listener's conviction that *human gesture* is

in some way responsible for the genesis of the sound. The concept of *gestural surrogacy* put forward by Denis Smalley (1986, 1996, 1997) helps to describe how we listen in this way. Four levels of *gestural association* (decreasing in associative potency) are identified: 1st-level: source identification (for example, violin); 2nd-level: identification of a group of likely sources (for example, unspecified stringed instrument); 3rd-level: gesture inferred ('where a gesture is inferred or imagined') through consistencies in sonic qualities (such as duration and attack-sustain-release (ASR) curves) that inspire gestural origin; and 4th-level: remote surrogacy (where 'source and cause become unknown and unknowable as any human action behind the sound disappears' Smalley 1997: 112). Levels 3 and 4 (particularly) appear to offer a way of understanding abstract musical materials - that is, gestural associations (real or imagined) tend to play a major role in the appreciation of sounds whose source/cause is unclear or unknown. Daniell Barreiro (2010) explains the significance of performative associations within the context of 'heavily processed sounds' as follows:

> These considerations could raise questions regarding the kinds of sound–gesture connections that may take place in the case of heavily processed sounds that do not seem to be linked to known gestures as their causes. Such sounds – that could be understood as examples of third-order surrogacy, using Smalley's terminology – may well be just appreciated in terms of spectromorphology (focused, therefore, on their intrinsic aspects). On the other hand, they could also be related to imagined causes by the listener, in which case it would be interesting to understand the main associative reason that supports such a relationship. (Barreiro 2010: 40)

This important concept, concerning both imaginary and real gesture-to-sound relationships is examined further in *Chapter Seven: Gestural Associations.*

Environmental Connections

As discussed in chapter one, the relationship of *the human body to the natural environment* logically underpins an understanding of the listening experience - placing it within the context of the soundscape of the natural

world. Further evidence supporting this important claim is presented in the second half of this chapter. The relationship highlights the likelihood of entanglement between all modes of listening – a subtopic that will be discussed throughout the book through inclusion of references to the environment in almost every chapter. The environmental mode of listening will also be examined independently in both wide and narrow contexts in *Chapter Eight: Environmental Connections.*

Spatial Forms

The way in which humans tend to create abstract spatial forms as a means of reducing, encapsulating and understanding the wealth of information received through the senses forms the basis of this multisensory mode of listening. It is also one that relates to the apprehension and understanding of structure and large-scale form. This important mode of auditory attention, concerning both imaginary and *real* spatial forms (encapsulating dimensions, shapes, territories, and relationships between spaces) is examined at depth in *Chapter Nine: Spatial Forms.*

Structural Concepts

The focus of discussion to date has concerned *the sonic object* or *event* but in the context of sound-based music, it is also important to consider the larger timeframes of sound organisation. It is here that *form* (musical structures that contains multiple sonic events) begins to emerge. Structural layers are produced by *inductive* (in-flow) and *deductive* (out-of-flow) listening strategies. These two divisions also tend to correspond to specific timeframes that house *component structures,* and the general principle of *structural integration.* Focussing on this aspect of sonic information is often referred to as *structural listening.* This topic will be fully addressed along with numerous contributing principles in *Chapter Ten: Structural Concepts*.

Research From the Fields of Psychology and Neuroscience

There is a large body of evidence from the fields of psychology and neuroscience that not only supports the current proposal (that the

listening experience is inherently multimodal and multisensory) but also adds much detail to it. The following paragraphs present the main contributions to this topic, the most significant of which have been proposed fairly recently (after 2010).

The Cognitive Psychology of Human Audition

McAdams and Bigand's '*The Cognitive Psychology of Human Audition'* (1993) contains several suppositions from prominent psychologists concerning the complex nature of auditory perception (the modes of listening with which humans glean information from sound sources and the processes by which the brain separates and categorises various sonic characteristics). Bregman (1993) identifies three modes of listening, two of which are learnt and one of which is proposed as a true universal. He states: 'One is the activation of learned schemas in a purely automatic way … A second process that can decompose mixtures is the use of schemas in a voluntary way…. [and the third is] primitive audio scene analysis.' (Bregman 1993: 13).

Bregman's first mode is concerned with *automatic information-gathering* based on the listener's ability to 'recognise and respond' to learned schemas, a skill that is necessary for safe navigation using the auditory sense as a warning system. The second mode concerns the human ability to *activate and focus attention* on one or more sound sources that contains learned schemas (message-conveying, abstract-musical, etc). This voluntary ability – also known as the cocktail party effect (allowing a single conversation to be deciphered in a room full of conversation) results in a deeper conscious probing of sounds and the information they contain. The third mode – the *primitive listening mode* - relies on 'general properties' (acoustic regularities supplied by the environment). The key argument to sustain the proposal of this third mode was originally put forward by Shepard (1981):

> Because animals have evolved in a world that contains regularities, it is likely that their perceptual systems have evolved modes of operation that take advantage of them. That is, as animals evolved, the properties of their perceptual systems became tuned to the regularities of the physical world.

> The resulting match is referred to as 'psychophysical complementarity'. An example of a general property is the fact that many sounds of our environment are harmonic. Bregman argues that these universal properties 'can be used for decomposing all types of mixture.' (Shepard 1981: quoted by McAdams and Bigand 1993: 15)

Shepard's comment is highly significant, as it produces a paradigm shift in the perceived value of *the environment* and its impact on the listening process – effectively converting it from an unremarkable listing as one of Smalley's (1996) nine indicative fields, to an archetypal lineal relationship responsible for the evolutionary development of human audition. It redefines *the environment* together with its universal laws and constants as the foundation of the listening process. Bregman (1993: 17-28) was able to specify this relationship further through a series of simple observations he referred to as 'acoustic regularities.' He states:

> **Regularity 1:** Unrelated sounds seldom start or stop at exactly the same time.
>
> **Regularity 2:** Gradualness of Change
>
> - a single sound tends to change its properties smoothly and slowly
> - a sequence of sounds from the same source tends to change its properties slowly
> - as the auditory system continues to hear two populations of tones, segregation occurs.
>
> **Regularity 3:** When a body vibrates with a repetitive period, its vibrations give rise to an acoustic pattern in which the frequency components are multiples off a common fundamental.
>
> **Regularity 4:** Many changes that take place in an acoustic event will affect all the components of the resulting sound in the same way at the same time.

Fast and Slow Modes of Listening

Bregman, et al, (1993) deduced that the width and depth of sonic information is determined by:

> a) our tendency as human beings to listen and react to sounds in ways governed by the evolutionary development of our species and

b) a matrix of learned associations accumulated by individual listeners over their lifetimes through exposure to their environments.

These general categories also represent the beginning and end points of a continuum between relatively automatic (fast) and reflective (slower) modes of auditory attention, although it is likely that these two extremes work in parallel rather than in sequence (as will be shown later in the chapter). The primitive mode of listening is connoted as a relatively automatic response driven by genetic hardwiring. Due to the relatively simple features exhibited within this mode of listening it might naturally operate (or at least begin) in the early stages of listening. Mari Riess Jones and William Yee support this deduction:

> Issues surrounding attending to auditory events that emerge from current theories fall into two categories – those concerned with relatively fast events and lower-order motion properties, and those concerned with relatively slow events and determinants of sustained attending. Topics associated with fast and/or brief auditory patterns include streaming, harmonic/spectral grouping, and determinants of perceived temporal order … (see Bregman and Warren) The present chapter is confined to issues within the second category, namely those involving sustained attending to extended auditory patterns (Riess Jones and Yee 1993: 85).

Bob Snyder (2000) proposed an explicit sequence of events that occurs in the reception of musical materials (offered below in paraphrased form). His description adds to our discussion in that it highlights multiple areas within the human brain that all contribute to the listening process.

> Sound waves enter the outer ear, stimulating the 'auditory nerves' leading to the triggering of 'echoic memory', which functions to store sound fragments (usually in the order of 250ms). Features are then 'extracted' and 'perceptually bound' to form patterns. Patterns are then analysed by the short-term memory, which is able to process at least several sonic events simultaneously as well as 'replay' patterns in a process of 'rehearsal' for the purpose of comparison with known schemas stored in the long-term memory. Recognition occurs when similar schemas (up to several at a time)

> are recalled from the long-term memory resulting in the assignation of meaning (paraphrased from Snyder 2000: 19, 47, 53, 71, 82, 96).

Neuroscientist, Edwardo Miranda (2010), offers a description that embraces a simultaneous or parallel processing metaphor rather than a sequential stepwise description (as Snyder suggests). Within the context of the current discussion, it suggests that multiple, or *all* modes of listening are likely to be active simultaneously. He writes:

> Cortical neural processing of music is an incredibly complex affair, which is still not well understood. It is clear, however, that the brain employs hierarchical neural structures to process music (Griffiths, Bu, Frackowiak and Patterson 1998; Patterson, Upperkamp, Johnsrude and Griffiths 2002; Stewart et al. 2008), and these processes may not necessarily happen sequentially. The brain is a complex distributed processing system, with various structures operating concurrently and at different time scales, from short-term to long-term musical forms. For instance, whereas lower-level structures may take care of processing the pitches of a sound sequence, higher-level structures would take care of processing the melodic contour engendered by the pitches of those sounds. But these processes might not necessarily be bottom-up; higher-level structures in the brain may make estimations of how the contour should evolve and this may influence how lower-level structures process pitches. (Miranda 2010: 20)

Miranda maps different functions of musical analysis to different areas of the brain – specifying 'early' and 'cortical' (outer layer) pathways - which generally correspond to our earlier comments regarding fast and slow modes of listening. He writes:

> The auditory system may be divided into two pathways, or stages: early auditory pathways, which lead from the outer ear through a number of subcortical regions and terminate in the primary auditory cortex, and cortical auditory pathways, which pass out activation from the primary auditory cortex to a number of other cortical areas … Cortical pathways enjoy greater plasticity than early pathways and their functioning are subject to learning and acculturation. However, the fact that sounds go through the same early pathways does not mean that everybody processes them in the same way at this stage. The brain differs slightly from person to

> person; therefore, degrees of functionality may differ slightly in the early auditory pathways. (Miranda 2010: 14)

Miranda states that although research from the field of neuroscience has so far been limited to address only traditional musical materials (pitch, rhythm, form, etc), it is expected that 'a number of facts about music processing in the brain… are likely to apply across the board, despite the existence of a plethora of musical genres and different philosophical takes on what music might or might not be.' (Miranda 2010: 14). His model appears below in figure 2.1:

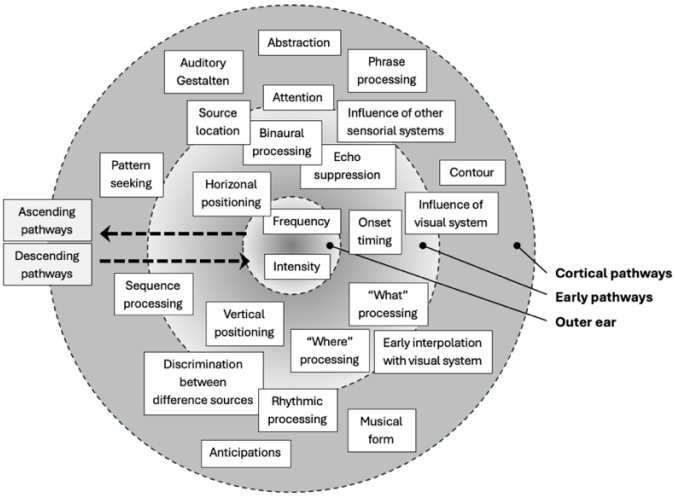


Fig. 2.1 Tentative functional map of the musical brain (Miranda 2010: 14).

Gary Kendall (2010) has also written extensively on the subject of multimodal listening with reference to the *sonic event.* Evidence from the fields of psychology and neuroscience support the view that *event schemas* act as psychological frames that house both sounds (sonic events) and their associative qualities.

> While an event is active, there will certainly be numerous other items that are active within the listener's mental spaces. Among those items will be ones that the listener will have reason to bind with the event. After the event has concluded, access to this non-dynamic event in memory will be limited to those items that were bound to the event within the active mental space. (Kendall 2010a: 66)

Kendall (2016) expands on this supposition by adapting a model of 'five mental layers' from Aage Brandt as a means of highlighting the 'multilayered' nature of listening. Of particular interest is the way in which he incorporates schema theory into his model. He includes the following reduction (from Brandt) in his initial argument:

> Our organization of what we call *meaning* is a process that occurs on many levels simultaneously. On a scale extending from the most dense, massive, and 'concrete' sensations to the most transparent and 'abstract' notional conceptualizations, we may distinguish a series of strata interrelated by processes of integration. (Kendall 2016: 33 referencing Brandt 2006: 173)

He then presents 'five mental layers' that he considers to be active in the process of listening to music and sonic art.

- Layer 1. **Sensations** – Perceptual organization and constancy of immediate sensations.
- Layer 2. **Gist** – Framework of things and space extended over several seconds enabling sustained awareness in the short term.
- Layer 3. **Locus** – Self-governing action in response to situations in the 'perceptual present' and slightly beyond.
- Layer 4. **Contexts –** Framework for enlisting and assessing medium and long-term, event-orientated schemas and expectations over an extended time frame.
- Layer 5. **Domains** – Frameworks of background knowledge providing long-term constancy. (Kendall 2016: 33)

Regarding each of the five layers, Kendall refers to *sensations* as 'the simplest' because it produces a 'sensory hold' – affirming the notion of the 'auditory present' (a duration of around 250ms). All other layers have

relationships with 'working memory' (of differing timeframes) and 'long-term memory' (schematic associations). It can be said (generally) that the layers of the listening process move from short to longer timeframes of operation, rather than occurring as a 'staged' process. The term '*gist'* has been carefully chosen as it highlights the mental layer in which the 'point or substance' (Collins Dictionary 2025) first becomes apparent. Significantly, there seems to be information passing both ways between the layers, as listeners check for confirmation (between layers) before assigning meaning to the sonic event. Individual layers may of course produce different associations leading to split or incommensurate meanings being produced. Kendall states that we seek out 'harmonisation' between all five layers, and that this often occurs naturally in real-world events. However, in art, various components are often incommensurate. Rather than this being thought of as a flaw, Kendall claims that it is often this mismatch that provides artistic interest (2016: 42-43). It seems that this type of incongruity is often experienced and enjoyed in art, as *attention* may be contemplated separately from *focus*, and focus may be *directed* to specific layers of mental activity. This explanation offers some similarity to Bregman's involuntary (attention) and voluntary (focus) modes of *decomposing auditory mixtures.* Listeners also tend to repeat established listening patterns (primarily obtained from previous listening experiences), so these three steering concepts (attention, focus, and repeated listening patterns) are always being playing out in standard listening experiences (Kendall 2016: 37-39). A reconstructed diagram showing the proposed connections between all five mental layers is presented in figure 2.2:

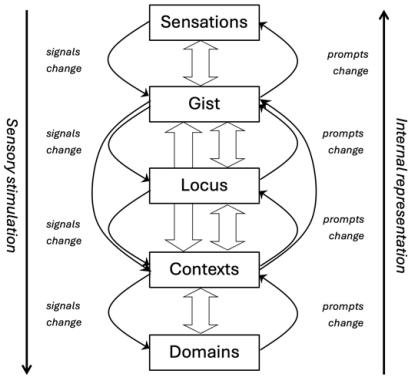


Fig. 2.2 Harmonization within the five-layer model (Kendall 2014: 196)

An Overlay of All Modes and Principles

It can be deduced (from this synthesis of research) that various modes of listening may be placed along a continuum ranging from fast, relatively automatic responses to slower, voluntary responses. Miranda (2010) has termed these areas 'early pathways' and 'cortical pathways.' The two categories respectively represent the end points on a fast-slow continuum corresponding to multiple processing areas within the brain. More sustained and cognitively focussed modes of listening include the information-gathering mode, modes concerning the recognition of complex musical codes, and may also include modes related to complex emotional responses (this topic will be discussed further in *Chapter Three: Musical Expression*). A vitally important supposition is that modes do not necessarily build on one another sequentially but appear to develop simultaneously (in parallel) in different areas within the brain with crosstalk occurring between all processing centres.

Early pathways are the timeframes from within which fast intuitive reactions to sound occur. These early auditory processes might include Kendall's mental layers of *sensations* and *gist,* Bregman's *primitive schemas*, Schaeffer's *mode two*, and Lennox and Myatt's *ambience labels*. Smalley's (1996) *reflexive* relationship might also be included (producing

basic 'feelings of pleasure and displeasure'). An example of a sound processed by *early pathways* might be a suddenly encountered fast moving vehicle (perceived as rushing' towards the listener). This type of sonic event tends to capture attention quickly and easily and often results in immediate emotional responses. Although sensory perception of this type is discussed as *near-universal* in healthy humans, it is important to register, as Miranda (2010) has, that humans are not equipped with identical audio-sensory apparatus. There is evidence to suggest that some listeners may *hear* in different ways. Miranda also notes that 'the brain differs slightly from person to person; therefore, degrees of functionality may differ slightly in the early auditory pathways.' (2010: 14).

Cortical pathways encompass the late perceptual stages of auditory attention. A network of modes of listening combine and cross-reference one another to produce much of the complexity of the total listening experience. These concepts and modes might include Kendall's mental layers of *contexts* and *domains,* Smalley's *interactive* and *indicative* relationships, multisensory associations, such as *visual*, *gestural*, *environmental*, *spatial*, and *structural modes of listening*. Again, it is important to iterate that these modes do not tend to develop sequentially, but rather simultaneously. It must also be acknowledged that differing cultural backgrounds and life experiences prepare the ears for differing interpretations and subjective engagement with musical materials. Naturally, these differences directly impact on different associations being made. While a composer may have certain understandings and a deep cultural relationship with specific materials and ideas, there is no assurance that the musical codes, or references employed will be universally understood. It is here that individual life experiences play a large role in establishing meaning (of the various sorts discussed) - as Suk-Jun Kim (2010) explains:

> Listeners engage with electroacoustic music by imagining a relational and developmental process among the sounds they hear. In fact, the use in many electroacoustic compositions of everyday sounds (and the characteristics we associate with them) naturally guides listeners into a

> process of imagining relations. The everydayness of the sounds emboldens them to 'picture' what they are hearing as they take cues from the way the sounds sit, move, integrate, disintegrate, or generally relate to one other. If image is 'the mode of givenness pertaining to the total imaginative presentation' (Casey 1976: 55), then sound-image can be understood as a mode of givenness initiated and led primarily by aural presentation that, nonetheless, constructs the totality of the imaginative presentation through the act of imagining. (Kim 2010: 46)

The following diagram (figure 2.3) overlays the key modal listening and auditory processing models discussed in chapters one and two, positioning them along an auditory processing time continuum.

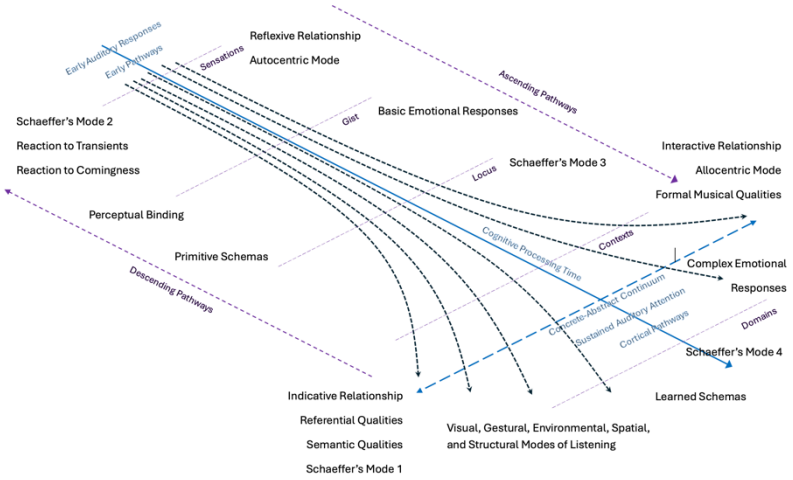


Fig. 2.3: An Overlay of Modal Listening and Auditory Processing Models

An array of compositional techniques and creative potentials emerge from this model. Sounds attended through early pathways - for example, those perceived as *rushing* towards the listener or *striking* in their transient qualities (corresponding to *sensations*) - will generally capture attention easily and may be *used* to deliver certain important types of

information. To promote a *reflexive* (emotional) response in the early stages of listening – operating at the mental layer of *gist* - a composer may attempt to make use of an acoustic regularity (for example, harmonic/spectral groupings). As auditory attention deepens, it is likely that most listeners will begin to separate out information-carrying and abstract musical characteristics. *Visual, gestural, environmental,* and *spatial* associations may also begin to develop certain meanings through the subjective responses of listeners, (these will differ somewhat between individuals). Such sounds may be appreciated as operating at the mental layers of *contexts* and *domains.* Musical relationships may also begin to form around materials that are perceived as containing *formal* musical qualities such as pitch and/or rhythm. Using Smalley's terms, if relationships between the early *reflexive* (basic emotional), and later *interactive* (formal musical) modes are retained, then this might give rise to the emergence of more complex emotional responses (which may also develop in response to emerging *indicative* (referential) relationships). Expectations of development and *structure* (both musical and dramatological/narrative) may also come to fruition together with a motivation to identify relationships between all constituent parts (towards structural integration). These processes are, generally, managed by the *outer cortical pathways*. Each listener's response will be unique to their own needs and the set of contributors that govern their listening experiences; however, various common reactions may also play a role in steering individual responses. Within this multimodal network, it is expected that there will be *crosstalk* occurring between the various areas of the brain responsible for processing the multiple modes of listening. Given the *parallel processing* power of the short-term memory, it can also be deduced that at least several (and possibly all) modes of listening will be in play simultaneously during the ordinary listening experience. Distinctions between attention and focus allow listeners to *tune into* the five mental layers of experience that operate concurrently.

Summary and Conclusion

While the visual and conceptual overlay of the numerous modes of listening presented in figure 2.3 reconciles most discrepancies between

established models, it does not capture the multiple levels of conceptual *nesting* evident in each mode of listening, or the nature of entanglement between them. In Smalley's system of categorisation, three main relationship types are specified (reflexive, interactive, and indicative) and within the indicative relationship alone, nine sub-fields are identified (gesture, utterance, behaviour, object/substance energy and motion, environment, vision, and space) - after which there are further divisions of concepts within each field (for example, the four levels of gestural surrogacy). A more straight forward and contemporary approach might simply consider a *network* of contributing listening modes and imagine crosstalk occurring between them. This way of thinking encapsulates all concepts and modes of listening discussed into a defined set of constituents - emotions, feelings, and expressions, formal musical qualities, referential meanings, real and imagined visual images, gestural associations, environmental connections, spatial forms, and structural concepts - that combine to form the totality of meaning from the sonic event in the minds of listeners. To simplify the categorisation and terminology of the many contributing fields identified, just eight interwoven modes of listening will now be put forward. These include: the expressive, formal, referential, visual, gestural, environmental, spatial, and structural qualities of sound, which combine to present a network of entangled listening modes. To aid clarity, these relationships are listed in table 2.1.

Listening Mode	*Sound Qualities*	*Listener's Reactions*
Expressive	Nuances	Emotional responses
Formal	Abstract patterns	Musical decoding
Referential	Referential and/or semantic	Assignment of meaning
Visual	Visual or quasi-visual	Association with visual images
Gestural	Performative movements and/or vocalisations (inc. instrumental)	Bodily and human associations
Environmental	Environmental patterns	Visual/Spatial associations
Space	Spatial forms	Abstract spatial constructs
Structural	Numerous patterns/concepts	Closure, insight

Table 2.1: Eight Modes of Listening

While not claiming to be comprehensive, the model relies on a network of concurrently operating modes of listening, meaning that the same materials/qualities may be viewed from a variety of different perspectives. Traditional divisions such as Smalley's tripartite model or the nine indicative fields may operate freely within it, and non-listed listening strategies (such as Smalley's *technological mode of listening*) may be freely added. Most importantly, the model naturally encourages all sonic qualities to be examined as an entangled network. For example, Smalley's *energy and motion fields*, may be examined from multiple perspectives including: the impetus behind a musical expression, a formal musical quality or pattern, gestural qualities perceived to be produced by human agents or performers, and/or the size and shape characteristics of abstract spatial forms. It is important to note that listeners tend to develop such associations simultaneously.

These relationships tend to be steered and reconciled by crosstalk occurring in multiple parts of the brain corresponding to early and sustained processes of auditory attention. Each mode of listening may be further conceptualised as a continuum ranging from implied to explicit levels of information corresponding to the five layers of mental activity: sensations, gist, locus, contexts, and domains. The simultaneous operation of all eight modes of listening (expressive, formal, referential, gestural, environmental, spatial, and structural) may therefore be conceived as eight different types of information (ranging from non-specific to specific) in combination, with the overall meaning of the sound steered by the *strength* of the constituting modes of listening. The nature of relationships between the eight modes of listening identified may also be conceptualised as a network diagram. A graphical representation of this principle is presented in figure 2.4.

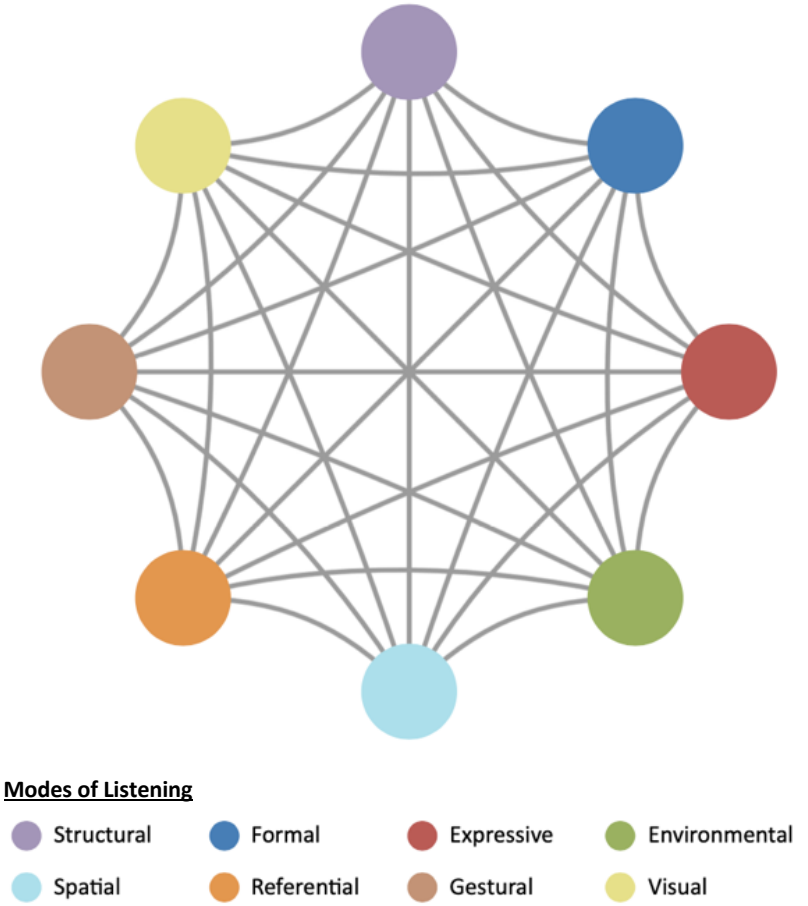


Figure 2.4: Eight Entangled Modes of Listening

The eight-way network diagram offers a useful way to conceptualise all possible connections and groupings that may occur between its contituent parts. Much can be gleaned from the model including voluntary focus on specfic modes of listening, two-way, three-way (and up to eight-way) modal combinations producing signature patterns and

groupings - some of which tend to be more common than others. All possibilities (in terms of modal wighting and combination) exist at the same time, with focus and attention freely shifting between groups and subsets automatically and/or voluntarily. These shifts may be further conceptualised as focus and attention *pathways* unique to each individual listener, or as *signature listening strategies* that are often replayed during repeated listening experiences.

These key concepts will now be examined at depth with the aid of numerous examples drawn from a variety of artmaking contexts. The subsequent chapters (three to ten respectively) examine each of the eight modes of listening both independently and as a network of contributors supplied through the ordinary listening experience. The purpose of this deep-level inquiry is to help gain a better understanding of their multimodal functionality and their application to the creative work of sonic artists and composers. As will be shown, the model supplies higher levels of scrutiny and understanding of the functionality of sonic materials through the multiple perspectives it offers. However, this should not be seen as a departure from previous (constituting) models, but as a framework for considering them all at the same time and in relation to one another.

For the purposes of our discussion, this principle has been termed *Sonic Entanglement,* which, by way of definition, describes the dynamic interplay of concurrent listening modes that emerge through the creation and reception of sound-based music.

Chapter Three: Musical Expression

Emotion and *feeling* are core to the musical listening experience. While we often use these terms as equivalents, they function differently. The following definition provides clarification.

> Despite the words being used interchangeably, *emotions* and *feelings* are actually two different but connected phenomena. Emotions originate as sensations in the body. Feelings are influenced by our emotions but are generated from our mental thoughts. (Psychology Today website – accessed 06:06:25)

The key distinction between the two terms may be distilled to the *subconscious* nature of emotions and the *conscious* nature of feelings. Based on previous deductions, basic *emotions* may therefore be closely aligned with auditory processes that are carried out relatively automatically, without deep reflection (through the early pathways of auditory attention), distinct from *feelings* that may be more generally associated with sustained auditory attention and/or contemplation (through the cortical pathways of the auditory attention).

The term *expression* (according to the Oxford English Dictionary 3rd Ed.) is a much broader one, encompassing at least emotion, feeling, passion, intensity, poignancy, intonation, tone, nuance, artistry, spirit, imagination, vividness, power, and force. The way it is used in the title of the chapter (*Musical Expression*) suggests reception of, or intention to convey, one or more of these sentiments. However, before this main topic of the chapter can be addressed, we might first consider the range of emotions and feelings that are commonly experienced through music.

Emotional Responses to Music and Sound

Denis Smalley proposed the *reflexive relationship* as one that contains 'positive/negative emotional reactions to sounds' – stating that 'as far as listening to music is concerned the reflexive relationship popularly dominates.' (1996: 80, 82). It has been observed that this basic type of emotional response tends to form in the early stages of the listening

process – alongside other important survival-related reactions such as: Schaeffer's mode two (sounds that impinge on our consciousness), Lennox, Vaughan and Myatt's ambience labels (mid-high frequency components), Bregman's primitive schemas (acoustic regularities) and all other aspects of early auditory attention right up to Schaeffer's mode three (the appreciation of musical patterns). Smalley suggests that sensory-led reactions and early forms of pattern recognition play major roles in forming the reflexive relationship. Interestingly, this deduction, when placed alongside those made in chapter two (with reference to. *early pathways*) suggests that basic types of fast emotional responses may also have played a role in the survival of our species.

Within the context of listening to music, emotions are deeply embedded. Suzanne Langer makes an early - yet important - contribution to this topic when stating that music is often accompanied by a 'feeling of beauty' or 'aesthetic emotion'. To prove its existence, she attests to the presence of the 'same emotion' in the heart/mind of the producer. She states:

> It has often been stated that it is the same emotion which overtakes a mathematician as he constructs a convincing and elegant proof; and this is the beatitude which Spinoza, who knew it well, called 'the intellectual love of God'. Something like it is begotten in appreciation of art too, though not nearly in the same measure as in producing; but the fact that the difference is one of degree makes it plausible that the emotion springs from one activity which the artist and beholder share in unequal parts – the comprehension of an unspoken idea. In the artist this activity must be sustained, complete and intense; his intellectual excitement is often at fever pitch. The idea is his own, and if he loses his command of it, confused by the material or distracted by pressing irrelevancies, there is no symbol to hold it for him. His mind is apt to be furiously active while an artistic conception takes shape. To the beholder the work is offered as a constant source of an insight he attains gradually, more or less clearly, perhaps never in logical completeness; and although his mental experiences also wakens the characteristic emotion, variously called 'feeling of beauty' aesthetic emotion and aesthetic pleasure, he knows nothing like the exhilaration and tense excitement of an artist before his pristine marble or clay, his unmarked canvas or paper, as the new work dawns in his brain. (Langer 1956: 210-211)

Langer claims that this 'depth of excitement' is formed through 'sustained activity' on the part of the artmaker. Her description refers to different levels of 'aesthetic emotion' in the receiver and the producer, with the latter displaying a higher degree of the same *feeling* that is derived from the listening experience alone. Her description of the producer's feeling as 'sustained, complete and intense' also adds to our comments concerning early and sustained modes of auditory attention (note her use of the term 'feeling'). Other researchers within the field of psychology, such as Elina Packalén (2008) present the possibility of more complex forms of communication through musical symbolism and the presentation of 'vital forms' to listeners. According to Packalén, these reactions are based on emotional responses to sound but are also open to subjective associations with more general aspects of life:

> The forms of feeling communicated by art are thus "vital forms" that are can't to human beings because they are analogues of life as psychically experienced. To consider music a kind of symbol is, however, not unproblematic. An object is used as a symbol, for instance a ring as a symbol for commitment, we do not pay so much attention to the symbolizing object, but to what it stands for. In contrast to these kinds of symbols, music seems to be the focus attention in itself, not in an instrumental way because of what it symbolizes. Nevertheless, Addis suggests that "to be an artwork just is to be the object or event that both symbolizes something and that is of interest itself. In order to understand Addis's theory of how music symbolizes emotions and states of consciousness, we must penetrate his classification of three kinds of representation. In general, in Addis's theory a representation is anything that something other than itself to mind, in either conscious or unconsciousness. An example of conventional representation is language where connection between a symbol (for example, the word "red") and its referent objects is not based on their inherent relation, on resemblance, or natures of the symbol and its user. Instead, the characteristic feature of tonal representation is that, in principle, the symbol that represents object according to a convention can be changed by choice. By contrast, in natural representation there is no such possibility of by choice, since a symbol that represents something naturally determines virtue of its inherent nature that it is doing the representing and what it represents. (Packalén 2008: 47)

In presenting the listener's *choice* as a functional component of musical discourse, Packalén (and Addis) extend on Langer's views with the notion of a shared *aesthetic emotion* resulting in a range of vital forms related to the listeners' subjective life experiences. While Packalén's work largely concerns *symbolism* (a term drawn from linguistics), it also presents a broad paradigm encompassing both *emotions* and *feelings* in musical contexts.

Expanding on Packalén's relationship with 'life experiences', a more contemporary view is that body-centred emotional categories may be 'linked to certain archetypical circumstances' (Kendall 2014: 193) leading to a range of subjective feelings in listeners. It is thought that these contextual circumstances - drawn from life experience - might be shared across cultures. In a recent study conducted by Berkley scientists (Cowen, Fang, Sauter and Keltner 2020) involving 2500 participants, it was found that the subjective musical experience across cultures could in fact be mapped to *thirteen* overarching feelings: amusement, joy, eroticism, beauty, relaxation, sadness, dreaminess, triumph, anxiety, scariness, annoyance, defiance, and feeling pumped up. This research supports the notion that emotional reactions to music tend to coalesce in common subsets of feelings. However, this produces yet another question: if the experience of listening to music evokes a subset of feelings, what is the full range of emotions and feelings that may be experienced by human beings? Allan Cowen (2025) answers this question directly in attesting to the existence of twenty-five distinct categories of emotion bridged by continuous gradients. He lists these as:

> admiration, adoration, appreciation of beauty, amusement, anger, anxiety, awe, awkwardness, boredom, calmness, confusion, craving, disgust, empathic pain, entrancement, excitement, fear, horror, interest, joy, nostalgia, relief, sadness, satisfaction, and surprise.' (Cowen, 2025: 1)

In comparing the results of these two surveys, the experience of listening to music appears to commonly trigger around 50% of the full range of emotions and feelings commonly experienced in everyday life. Having said that, these inventories of emotional categories rely on words - terms

that may be interpreted in several ways – rending this type of reductionist thinking inaccurate. Other researchers (Plutchik, Deacon, et al.) have attempted to identify an even smaller number of emotional categories linked with 'archetypical circumstances' and have found evidence that complex emotions are in fact produced through combinations of more basic emotions. Gary Kendall explains this *building-block* principle:

> One way of resolving all this apparent inconsistency is offered by Robert Plutchik (1980), who proposes that many emotions for which we have words, and which are a part of social life are in fact blends of basic emotions. His idea is similar to a colour wheel in which there are primary mixtures of closely related emotions (joy + acceptance = friendliness) ranging to more distant combinations called secondary (sadness + anger = sullenness) and tertiary mixtures (anticipation + fear = anxiety). Terrence Deacon (2006: 38–9) echoes this idea and further suggests that mixtures of typically oppositional emotions produce 'emergent emotional states' that 'are transformed by their interactive relationship'. For example, 'awe may juxtapose both the joy of appreciation and the terrifying recognition of fragility'. That complex emotional states are, in fact, blends, touches on the broader topic of feeling blends to be addressed below. (Kendall 2014: 193)

Returning to Suzanne Langer's early research – in 1953 she puts forward an interesting proposition that these *patterns* (relating to a common *subset* of emotions) may be expressed through music, due to the 'similarity between the intrinsic structures of music and the structure of our emotive experiences' (paraphrased by Windsor 1995: 54). Through this construct, Langer proposes what she believes to be one of music's very special qualities – the preservation of musical *purity* (with reference to abstract musical qualities), together with the possibility of expressing a range of emotions that impact on listeners in archetypal ways. This proposition might not be such a leap, given Shepard's and Bregman's previous observations concerning *psychophysical complementarity* (human audition developing in response to frameworks presented by the natural environment). If Langer's points are valid, then regularities contained within specific types of musical materials might also present the framework that supports certain types of common emotions. It may be that such patterns are shared between nature, the auditory system,

and various categories of emotion. A direct example may be the *feeling* of tension and release (and emotional relief) in the human body at the point of musical cadence.

To illustrate the proposed emotion-to-feeling pathway, and to highlight the range of subjective responses that may spring from a single listening experience, I offer the following example. As part of my more recent experience, I have had the privilege of working with children. It has been interesting to observe the range of reactions and high levels of *feeling* produced in the minds of young people aged 6-8 (in a class of 20). One specific example concerns responses to an expressive solo performance (voice and guitar) of a well-known New Zealand folk song *Blue Smoke*. Reactions ranged from cathartic tears (some children became deeply upset) to blissful enjoyment of the melancholy, to expressive dance, to *resting* on the floor. Very few reacted with indifference – most appeared to be deeply affected in some way through a range of dynamic responses. There appeared to be a *trigger* - a pattern within the music, equating with an emotion (sad, nostalgic) that each child interpreted and reconciled in their own way through a range of subjective *feelings*. Bharucha and Curtis (2008) offer further clarification concerning these stated divisions in supporting the emotion-to-feelings pathway that has been under discussion. They claim:

> Musical feelings need not be about something, may or may not be valenced, and, unlike emotions (which are nameable, e.g., sad and happy), may not be readily nameable. Yet, they may have an affective quality in that they are felt and not just perceived. (Bharucha and Curtis 2008: 579)

While the multiple references we have used in the discussion to date have been useful in shedding light on the topic of emotional reactions to music and sound, they are defensible only to a point, as the majority do not clearly identify what kind of music (or sound) is under scrutiny. To highlight this limitation, I refer to Langer's comment below:

> Why then, is it [music] not a language of feeling...? Because its elements are not words - independent associative symbols with references fixed by

> convention. Only as an articulate form is it found to fit anything; and since there is no meaning assigned to any of its parts, it lacks one of the basic characteristics of language-fixed association, and there with a single unequivocal reference. (Langer 1953: 31)

Her comment clearly excludes the use of meaning-bearing materials (for example, words understood by listeners), that are commonly employed in a variety of contemporary music genres and categories. Her description may in fact be limited to an unstated personal preference towards classical instrumental music. This is extremely important to our discussion, as a modern approach to sonic artmaking does not limit the sound palette to a predefined instrument set. It encompasses *all* sound types. Luke Windsor exposes the narrow application of Langer's claims (encapsulated through the previous reference) in the following succinct statement:

> Acousmatic music challenges this ideology at its root. Here, musical units often consist of sounds which have pre-existing, 'referential' meanings. (Windsor 1995: 54)

His comment recontextualises the current discussion, from one that has so far gravitated towards abstract musical qualities and their ability to evoke specific emotions and feelings to one that embraces all sounds. This includes referential and semiotic qualities which, through their associative, dramatological, and narrative potentials, might logically extend to the full range of twenty-five emotions and feelings specified by Cowen (2025). That is, stories might be told to access more complex emotions, while formal musical qualities might be employed in their emotion-bearing capacity and ability to trigger the narrower range of thirteen commonly experienced feelings in listeners.

When used in combination, emotion-bearing *formal* and *referential* qualities (described by Smalley as 'the duality of sound') produce two combinations of listening modes that can be defined within the current context as *expressive-formal* and *expressive-referential*. In the case of formal musical materials, the *musical code* (a melody, chord, or other

pattern) is primarily responsible for delivering the emotional content. Smalley also observed this combination of modes in common use, stating that 'within the context of Western art music there should ideally be a partnership between the reflexive and the interactive' (Smalley 1996: 82). In the case of referential or semantic materials, the associative meaning (including dramaturgy and narrative) is normally responsible for delivering the emotional content (for example, a sad story). To be clear, the reference (for example, baby) may also supply emotion or feeling (for example, adoration) in a two-way modal blend that can be described as *expressive-referential.* Juslin and Sloboda extend on this latter principle, which they describe as *meaning* (derived from or steered by emotion). They state:

> We noted at the beginning that music and emotion is linked to the concept of meaning. … A review of this work suggests that musical expression is multi-dimensional: Throughout history, music has been regarded as expressive of motion, tension, human character, identity, religious faith, and social conditions, though the most common hypothesis is that listeners perceive music as expressive of *emotions* (Davies, 1994). In fact, music has been defined as "that one of the fine arts which is concerned with the combination of sounds with a view to beauty of form and the expression of emotions" (Oxford English Dictionary, 3rd ed.). (Juslin and Sloboda 2013: 588)

For Juslin and Sloboda, musical meaning is clearly based on our previously defined term '*expression',* which may be further broken down to the specified sub-fields of motion, tension, human character, identity, religious faith, social conditions, and (most commonly) emotions. This comment effectively expands our discussion to include various other common forms of expression beyond emotions and feelings. It also contrasts with Smalley's (1996) approach who recognises basic emotional responses (the reflexive relationship) but bases musical meaning primarily on *referential* associations (the indicative relationship) and the characteristics of the nine indicative fields gesture, utterance, behaviour, object/substance, energy, motion, environment, vision, and space. Again, the most likely cause for this split is the unspecified musical context - Juslin and Sloboda's proposition probably relates to instrumental music,

while it is known that Smalley's focusses on a multitude of sound-based materials featuring both referential and abstract qualities. Other contemporary views suggest less separation between expressive and referential meanings. Gary Kendall explains how this more complex and entangled type of meaning and understanding (drawn from everyday life) commonly develops within the context of listening to sound-based music.

> Our emotional state embodies our assessment of our situation, moment-to-moment and ongoing. Emotions are part of our cognitive engagement with the world, an intrinsic component of meaning and understanding. Our ongoing projection of outcomes affects us whether the context is the contingencies of everyday life or the evolution of an electroacoustic work. Certainly, we are aware of the differences between artistic and practical outcomes, but listening, even artistic listening, gives rise to meaning as a direct manifestation of this everyday cognitive process of finding meaning in the world around us. We can recognise a commonality between the ways that meaning in everyday life and artistic experience are embodied in feelings and emotions. We may differ about how to express our feelings about a particular passage of an artistic work and its meaningfulness to us, but we are all probably in agreement that it has some form of meaning of this kind. (Kendall 2010a: 65)

The fact that meaning can be commonly experienced as emotional (expression-based) and/or literal (association-based) is evidence enough of the level of entanglement between the stated modes of listening expressive, formal, and referential. Given the observed dominance of these multimodal combinations and the parallel processing power of the human brain (previously discussed as *layers of mental activity*), it is also likely that sophisticated *feelings* may develop through these two or three-way modal partnerships. Furthermore, it is likely that crosstalk, reflection, and multisensory modes of listening (visual, gestural, environmental, spatial, and structural) - may supply additional emotional triggers that may add to existing feelings or even create *new* emotion-to-feeling pathways. Kendall (2014) resists the notion of compartmentalised feelings/emotions and other modes of listening altogether, proposing instead a *blend* with all aspects of auditory attention. He states:

> Equally distorting is the notion that we can disassociate feeling and emotion from the act of perception. Perception is too often treated as a recording system for physical events, albeit an imperfect one. The objectivity claimed by the physical-acoustic perspective can thus be stretched to embrace a translation from physics to psychophysics. But, in the actual flow of human mental activity, perceiving and feeling are simultaneous, co-mingled, and fused in the experience of the moment. (Kendall 2014: 192)

The way in which emotions and feelings 'co-mingle' with each of the previously discussed five layers of mental activity presents a compelling subtopic. Feelings tend to begin to develop at the level of sensations where *qualia* – ('raw sensations of colour, pitch, warmth, etc as well as more dynamic sensations associated with the temporal flow of experience' Kendall 2014: 194) - come together to form patterns that are very quickly reconciled with various types of 'image schemas' (discussed in *Chapter Four: Formal Musical Qualities*) and flow dynamics ('a sense of the texture and energy flow that could be captured as rough, bumpy, grainy, smooth or flowing, etc.' ibid: 194). These early stages (*sensations* and *gist*) seem to equate with basic emotional responses such as feelings of like or dislike. Qualia and flow dynamics are naturally integrated with *all* modes of auditory attention; however, one such common combination is the previously stated partnership between expressive and formal modes (also highlighted by Smalley). Kendall (2014) provides a summary of how feelings together with musical patterns might be appreciated within the context of the *gist* and *locus layers of mental activity* discussed in chapter two. He states:

> Importantly, these are felt as well as understood as patterns. For example, the image schema Removal-Of-Restraint is understood as an embodied pattern of tension and release interacting with physical objects. Once the pattern is codified as a schema, it can be associated with events that persist beyond the normal time frame of physical gesture. So, for example, a section of music might be understood in terms of Removal-Of-Restraint and still retain its sense of physical tension and release even if it is minutes long. (Kendall 2014: 195)

Referential and/or semantic materials can powerfully present *expressive-referential* qualities, including dramaturgy and narrative. The assignment of emotionally imbued referential meaning tends to develop in the layers referred to as *locus* and *contexts*. The sound of a baby crying (for example) might not just deliver referential meaning, it might also stir various emotions leading to a range of subjective *feelings* in the listener. These feelings are stored within long-term memory schemas that carry referential information. Words like *dramaturgy* (describing the theory and practice of *dramatic* composition) are particularly useful in this context as they combine elements of emotion and narrative that are drawn from both contributing types of meaning (expression-based and reference-based). A reconstruction of Kendall's table presenting the differentiation of the five layers with respect to feelings is presented below.

Layer	Feelings
1. Sensations	Qualia, flow dynamics
2. Gist	Qualia, flow dynamics, short-term image schemas
3. Locus	Medium-term image schemas, schematic associations, certainty/uncertainty
4. Contexts	Anticipation, basic emotions, long-term image schemas
5. Domains	Consistency/inconsistency

Table 3.1: The differentiation of the five layers with respect to feelings (Kendall 2014: 196)

A brief summary of *emotional responses to music and sound* with reference to table 3.1 will now be offered. The mental layer of *sensations* includes the *expressive* mode of listening in its guise as a set of early unspecified emotional triggers. Basic musical patterns (*formal* musical qualities), imbued with impressions of gesture/behaviour (*gestural* associations), and motion (*spatial* forms) begin to become apparent at the mental layer of *gist* enabling the emergence of emotion-to-feeling pathways. Longer term schematic associations housed within the layer of *locus* are necessary to appreciate musical constructs such as tension and release, which function precisely through bodily sensations and feelings.

Referential qualities (associative meanings) also tend to be partially supplied by these schemas that operate within the layers of *locus* and *context,* with the latter accessing long-term image schemas. Emotions may be triggered within any part of this process leading to a range of subjective feelings. Predictably, modes and relationships related to *structural* concepts tend to be located within *contexts* and *domains* - layers that correspond to sustained auditory attention and associations with schemas stored in long-term memory (which also house *feelings*). Our developed understanding of how all the stated modes of listening operate is thus one of simultaneous interdependence heavily entangled with emotions and feelings.

Expressive Artmaking Strategies

As noted at the beginning of the chapter, the term *expressive* in the context of this discussion embodies a deliberate attempt on the part of the artmaker to *convey* specific sentiments to their audience. These expressive messages include emotional categories but also embrace more remote concepts such as 'motion, tension, human character, identity, religious faith, and social conditions.' (Juslin and Sloboda 2013: 588). A topic central to this chapter concerns the ways in which musicians (composers, performers, producers, etc.) might *convey* these expressive qualities. Although the exact nature of expressive communication may be indefinite, the *need* to express ideas with commitment has been a dominating tenet of most conservatoriums of music throughout the world. To investigate this important principle, we will now examine the subtopic of *feelings as a source of intention for musical expression.* As a starting point, I offer the following extract sourced from Encyclopaedia Britannica in an article entitled *Art as Expression* by John Hospers (2022).

> One view of emotional expression in art is that it is preceded by a perturbation or excitement from a vague cause about which the artist is uncertain and therefore anxious. The artist then proceeds to express feelings and ideas in words or paint or stone or the like, clarifying them and achieving a release of tension. The point of this theory seems to be that artist, having been perturbed at the inarticulateness of their "ideas," now feel relieved because they have "expressed what they wanted to express."

> This phenomenon, indeed, a familiar one (for everyone has felt relieved when a job is done), must still be examined for its relevance. … The idea is that the 'feeling' of intensity (see definition) with which the artist feels the need to express themselves might also be received by the audience. This seriousness and 'need' (not faked) might be seen to 'come through' as a message carried by materials during the listening process. (Hospers 2022: 1)

From this description it can be clearly seen how (in theory) a composer's or performer's *seriousness of intention* may be directly experienced as a *feeling* by the listener. It appears that if the artist is fully committed to an idea, their 'seriousness and need' may be translated to an audience. Thus, the communicability of the music has a direct relationship with the producer's intentions, placing a weight of responsibility on the composer/performer. What do they have to say? What is it they feel passionately about? These '*poietic* intentions' as discussed in chapter one (see Nattiez, 1990), form an important aspect of the communicability and potency of any given musical work. The current discussion investigates the *commitment* of the composer/performer as central to musical communicability. Such *serious ideas* are communicated to an audience primarily *through* emotional expression and reception of a feeling of *intensity*. The key argument is that music that does *not* contain such seriousness of intention can be *heard* to be deficient, less expressive and less communicative by comparison. Surprisingly, there seem to be research strands that support such a proposal - including the possibility that important information relating to intentions and emotions might be delivered through subtle *nuances* in the music itself. Bob Snyder (2000) provides one possible explanation concerning the non-quantifiable emotional qualities of musical nuance within the categories of pitch, rhythm and form. He writes:

> Nuance is the variation that takes place inside of the boundaries of a musical category. Musical nuance in this sense includes a range of different phenomena. Melodic nuances include the subtle bending of pitches, vibrato, and small variations in intonation. Rhythmic nuances include the rushing and dragging of beats and groups of beats, as well as small variations in overall tempo. Formal nuances can include repetitions of a

> pattern so subtle it is unclear whether anything about the pattern has actually been changed. Nuances are variations that we can notice on some level, but that do not affect our basic sense of which category we are hearing at a given time. They are represented at the level of echoic memory and are then usually lost in the later conceptual categorisation process. In other words, nuances may be present at the beginning of perceptual categorisation but are absent at the level of conceptual categorisation. They can be noticed, however, because they may be brought directly into perceptual awareness, bypassing the category structure of long-term memory. Thus, as we have seen before, we can experience more than we can remember. Although not remembered, nuances contribute to the subtle ongoing feel of the music and often carry important emotional information. (Snyder 2000: 86)

Snyder's statement not only provides an explanation concerning how 'emotional information' might be delivered through music but also supports the notion that the location of the *expressive* mode of listening has its genesis within the *sensations* and *gist* layers of mental activity.

The Concept of Agency as Expression

To gain a finer-grained understanding of expression as an artmaking strategy, the concept of *agency* must be included. Within the context of music and sonic art, the term *agency* is assigned a very specific meaning, referring directly to the *agent* (normally human) who is perceived as being responsible for creating the musical materials (note, sound, phrase, event). The concept of *agency* reconciles the vibrational energy of musical sound with the kinetic or *gestural* energy of the producer. *Agenc*y will be examined at depth later in the book (in *Chapter Seven: Gestural Associations*) with reference to Smalley (1996, 1997), Clarke and Davidson (1998), Kendall (2010a), et al, but for now the context of discussion will be limited to the topic of *agency as a vehicle for expression*.

Types of agents (actual and perceived) include at least: human, environmental, electrical, mechanical, and associations with all other forms of energy; however, the listening experience also tends to *anthropomorphise* non-human agents, naturally imbuing ontology, gesture, and expressivity to the musical materials produced.

Furthermore, musical agents are often perceived to be directly responsible for communicating the essential quality of *seriousness of intention* previously discussed. In this way, the agential release or communication of emotion is seen to constitute part of the energy structure (along with kinetic/gestural energy) that is converted to sound. The concept of *agency* always embodies an aspect of practical sound-making (for example, playing a musical instrument), even when the agents responsible for producing the sound are not physically present. For example, when *recordings* of human voices are used in music production contexts. Agents may be both real and imagined.

Agencies may also be expressed on a continuum. At one end, there is the gestural control and presence of the performer/producer in every musical decision, acting upon the materials to bring the musical work into creation. At the other end, the artmaker's agency may be psychologically reduced to make way for associations with *other* agents perceived to be contributing to the music-making process. Here, the artmaker may be seen to act as a *conduit* or channeler for such perceived agents. While composers/performers may be responsible for the literal act of sound-making they may not be perceived as the *sole agent(s)*. To examine this continuum, in the following paragraphs it will be separated into its two extremes, first discussing each distinct position, then presenting a mid-point synthesis of the two.

The Stance of Composer/Performer as Agent

This common music-making attitude refers to the overt agency and conscious decision-making of the composer and/or performer. In this stance, it is the feelings, concerns and emotions of the composer/performer *in-person* that are the focus of musical communicability. When the composer/performer is the sole agent there is a *seriousness of intention* that relates to the individual(s) and what they have to say: how they feel, and other aspects of their personalities drawn from human existence. In live performance, this is normally presented audiovisually through visual cues and nuance-level qualities within the music through a performer whose *personality* plays a major role in the

communicability of the work. Where the composer and performer are different people, ideally a partnership will be entered where both individuals work together to express the intention of the given work. To adopt this stance, a composer/performer might first consider what they have to say: who they are, how they feel, and what they have to discuss. It presents the possibility of eliciting emotionally centred ideas in subjectively meaningful ways, ideally resulting in the production/reception of authentic and convincing musical materials displaying *seriousness of intention*. Some of these feelings may have natural associations with life-events stored in long-term memory schemas and may also contain narratives (semantic meanings) that might form additional emotional, meaningful and/or musically functional messages. An example of this first stance in action might be to contemplate a loving relationship *while* expressively playing an acoustic guitar. In this act the composer/performer is attempting to imbue the music itself with their subjective emotions through *nuance-level* subtleties in articulation. Within the context of acousmatic music, we can imagine a scenario where the composer records and presents a deeply personal story together with accompanying sounds (for example, the birth of their child). Here, the semantic qualities of sounds (recorded speech together with the sound of a newborn baby crying) may automatically trigger a range of emotional responses. This method of *employing* sounds with emotionally affective qualities is direct, involving both story-telling and the activation of schemas relating to personal life-events stored in the long-term memories of listeners. The two examples highlight deductions already made in part one of this chapter – that both formal and referential materials commonly act as vehicles for personal expression.

The Stance of Composer/Performer as Conduit

We might imagine the contrasting position on the continuum as a *stance* where the composer/performer acts solely as a conduit for *agents* that are perceived as originating *from elsewhere.* Such *imaginary* agents are often recognised as residing outside the performance arena. For example, a classical performer may intentionally channel specific

emotions and feelings they perceive to be housed within the music itself. I refer to this stance as *composer/performer as conduit.* At this extreme end of the continuum, the aim is to reduce the perceived presence of the composer/performer to as close to nil as possible. The physical producer of the music ideally acts only as a channeler or conduit of expression in the service of intentions, feelings, or other messages that are perceived (by the producer and ideally the audience) as originating beyond the immediate orbit of the performance arena. Such is the approach of many Taonga Pūoro (indigenous instruments of Aotearoa) makers and players, who deliberately limit their personal input in the sound-making process and seek out non-present agents related to the material qualities of the instruments they are playing and ngā atua (the Gods) they are affiliated with. So deeply embedded and widely understood is this practice that Richard Nunns (2014) used the rather crass term 'cell phones to the divine' to describe the functionality of the entire instrument set. An atheistic view would suggest that is this simply a result of an active imagination, but there are those who would argue to the contrary. While it is beyond the remit of this book to enter such discussions, it is important to highlight the *dominance* of the approach in various music-making settings. Across cultures and timeframes, sound and music production has been in service to spiritual beliefs and practices.

While this subtopic of *composer/performer as conduit* presents some difficulties in terms of its defensibility, it is nevertheless, one that should not be shied away from. Over my thirty-year career as a lecturer in music, I observed *hundreds* of musicians engaged in related practices. Specific qualities that unite the range of approaches witnessed include a reduced sense of ego or self-importance, and maintenance of at least some degree of practical music-making centred on what is perceived as influence beyond the immediate self. Many of these professional musicians have expressed similar ideals - embodying an absence of physical and technical barriers to enhance this stance of *composer/performer as conduit*. Examples of music produced in this way are plentiful: the shofar bowlers of Jericho, shamans beating their drums, much of the music of Johann Sebastian Bach, Tibetan monks droning,

Gurdjieff de Hartmann and the Swirling Dervishes, the previously mentioned Taonga Pūoro players, John Coltrane's *A Love Supreme* and Terry Riley's *in C*.

When I questioned my peers (composers, performers, and musicologists) within the academic music fraternity on this seemingly esoteric topic (*performer/composer as conduit*), I expected to find mixed attitudes and even ridicule towards it. Instead, I found quite the opposite - a near-universal acceptance of, and complex private practices relating to the paradigm. Through these conversations I developed a new understanding of just how widespread this way of thinking and operating is. Concepts such as musician as conduit, connecting to the divine, bridging worlds, music therapy, personal transformation, emotional expression, worship music, and music for strengthening communities appear to be at the very *heart* of the music-making process and may even form an integral part of the zeitgeist of music in general. The fact that this topic is *unspeakable* in academic circles (with very few exceptions) does not reduce the leading role it plays in a staggering variety of music-making settings. Notably, decorated composers/sonic artists and academics Paul Rudi and Gary Kendall both speak openly about their perceived connections to the divine and specifically about the roles they play as *conduits* during the compositional process. Paul Rudi writes:

> I have been called "the High Priest of Sound," and my music described as "The Universe unfolding one sound at a time." The music that comes through me opens up a horizon that you may navigate on your own terms. The journey is not so much of my design, but of your desire. It will help you come home to yourself. Like the ceremonial spaces on my land, every path is unique and no one else's is like yours. I love helping people to more deeply experience life around them, and themselves as part of it! (Paul Rudi, sourced from https://www.paulrudy.net/about-paul-rudy 10:03:25)

Gary Kendall, (who has also been trained in Shamanic practices and energy healing), sees this as a way of linking life experience with divine notions of harmony and alignment:

> Practitioners of divination might explain this by saying that the state of making true statements is in alignment with divine truth and one can feel the harmony. The no-statements are out of alignment and one can feel the disharmony. There is every reason to believe that all human beings can experience these subtle shifts of being in or out of alignment with divine truth in everyday life. These shifts are gently guiding us toward living in alignment…This approach to compositional decision might seem similar to the chance systems used by John Cage. One of the ways those systems are typically understood is as a method that defeats conventional judgment. In that way my practice is similar to Cage. Another typical way of understanding Cage is that conventional judgment is defeated by randomness. The way I understand my practice is that it is convergent: I am attempting to uncover the right value for each instance, not a random one. (Kendall 2005: 3)

A Combination Stance: A Synthesis of Agents

The mid-point on the proposed continuum features a synthesis of agency in the form of local musicianship (the performer/composer), and non-present agents (perceived as originating from elsewhere). This is the basis of expression for most music and organised sound. To illustrate its various qualities, I use the example of Bach's *Goldberg Variations* played by Glenn Gould (1955), popularly described as *transcending cultural and temporal boundaries* due to its profound emotional content. Logically, a component of musical agency within the context of live performance may be assigned to the skills, artistry and *personality* of the performer. Further, following Nattiez (1990), other qualities might be assigned to be that of the 'poiesis' of the composer (posthumously) who may have been attempting to convey specific emotions or ideas to an audience.

The Goldberg Variations (J. S. Bach) played by Glenn Gould (1955).

Here, there is an obvious intention on the part of the performer to *channel* something that is beyond himself and his own performance abilities. Also evident is his *seriousness of intention* to use his highly developed skills as a musician to draw together a network of *agents* that might ultimately be perceived as the creators of the music – of which Gould becomes one. It is this powerful synthesis of Gould/Bach and beyond that creates such compelling musical discourse.

In the context of acousmatic music, this mid-point stance is also extremely common. The composer's poiesis plays an active role through the ways in which the sounds are *organised* and performed through loudspeakers, while recognisable recorded sounds (for example, an instrumental performer) provide strong links to the agents of *elsewhere.* Acousmatic music provides a very special vehicle in which the illusion of agential presence may be maintained through pre-recorded materials alone.

Although difficult to defend in academic terms, a minority of leading scholars from the field of electroacoustic music studies have highlighted the special ability of some music to present powerful emotions and feelings beyond the immediate orbit of the listener. In the following comment, Simon Emmerson directly supports the concept of the *mid-point stance* currently under discussion. He states:

> Western art music has thrived on the tension (even contradiction) between the work as transcendental, beyond individual interpretation, and the work existing only in its moment of performance, through an interpreter whose personality is certainly not "bracketed out". (Emmerson 2007: 29)

The comment is centred on exposing a *partnership* (between performer and external agent), where the performer does not normally act solely as a *conduit* for the spirit of the music but offers aspects of their own

personality to highlight specific musical qualities in cooperation with other perceived agents. Emmerson highlights this stance (*a synthesis of agents*) as the norm – where both present (performers/composers) and absent agents (non-present composers, recordings of human activity, or environmental energies) may be located both inside and outside the performance arena.

Historically, *Western art music* has made use of a fairly small subset of formal musical qualities such as melody, harmony, phraseology, and subtle dynamic control to express ideas associated with emotions, feelings, and the divine. Specific instrument sets too, such as strings and/or voice have tended to dominate. Samuel Barber's *Adagio for Strings* is an example of this approach – commonly producing strong emotions and feelings in listeners . Other composers have recognised the expressive potency of the combination of harmony and human voice. An example of this expressive synergy is Mozart's *Lacrimosa*. Of all the musical instruments, we are likely to notice the micro-inflections and subtle *nuances* of the voice the most, as we draw on deep archetypal relationships to human utterance. In some cultures, entire *genres* and categories of music have been created in the service of this expressive potency. Although it is beyond the scope of this chapter to provide a detailed overview of expressive compositional techniques within the context of instrumental and vocal music, it is important to note that this well-trodden pathway provides at least one method of effectively managing expression in sound-based music contexts.

Summary and Conclusions

In the first part of the chapter common emotional reactions to music and sound were discussed. Basic responses, initially relating to 'likes/dislikes' (Smalley 1996), appear to be formed in the early stages of auditory attention along with other modes of listening related to survival. Both composers, and researchers from the field of psychology have attempted to define common types of emotions and feelings associated with music. These include Langer's (1956) 'aesthetic emotion', Packalén's (2008) 'vital forms', and *thirteen* types of feelings describing 'subjective musical

experience across cultures' (Cowen, Fang, Sauter and Keltner 2020). Other researchers (Plutchik, Deacon, et al. 1980) present a developmental pathway from body-centred emotions to thought-based feelings. Langer (1953) puts forward that emotions and musical structures present *shared patterns* that may be both felt and perceived. From a more modern perspective, this paradigm might be evidenced through commonly experienced feelings of 'tension and release' in the human body coinciding with points of musical cadence (Kendall, 2014). Yet another approach proposed by Juslin and Sloboda (2013) is to consider expression ('motion, tension, human character, identity, religious faith, social conditions, [and most commonly] emotions') as *meaning*. This presents the wider view that all eight modes of listening identified (expressive, formal, referential, visual, gestural, environmental, spatial, and structural) are naturally imbued with emotions, feelings, and meanings. This discovery brings to the fore what Kendall (2014) refers to as 'the feeling blend' (the expressive mode of listening blended with all other modes of listening). Further evidence of these *blends* is observed through emotionally imbued early-to-late stages of the auditory attention process (pathways from basic emotions to more complex subjective feelings) with reference to the five layers of mental activity 'sensations, gist, locus, contexts, and domains' (Kendall 2014).

In part two of the chapter, *expression* as an artmaking strategy is discussed – that is, how a composer or performer might think and act to effectively manage expression in the music-making process. The concept of producing music with *seriousness of intention* is first examined, presenting the possibility that music featuring such qualities might generally be appreciated as more *expressive*. While this supposition appears to be reaching, a few researchers including Snyder (2000) support the claim under the proposal that information of this type is commonly delivered through *nuance-level* qualities within sounds. The concept of *agency* is then introduced to unite the range of perceived sound-producing energies from individual performers/composers who are directly responsible for the production of the sound to what Emmerson (2007) described as 'transcendental', and all that is between,

including non-present human agents, environmental and mechanical energies, and all other perceptions of influence from elsewhere. A continuum from *performer/composer as agent* to *performer/composer as conduit* is then examined, and a *mid-point* stance is identified where such non-present agents might be perceptibly integrated with or *channelled* through the composers' or performers' skills and personality.

From a practical sound-making point of view, this combination stance (*a synthesis of agents*) is historically relevant. Examples of music have been put forward as evidence of what is, likely, the most common historical context for the delivery of expressive musical content – *through* formal musical qualities (melody, harmony, phraseology, and dynamics) and nuance-level articulation. In acousmatic music, where prerecorded materials are employed, the pre-eminence of non-present agents is most striking. Here, where referential/semantic materials may also be employed, it is important to consider the emotional impact of associations, dramaturgy and/or narratives on the listener. This latter approach to triggering emotions and feelings relies on associations with the life experiences of the listener (housed by schemas stored in long-term memory), which carry with them emotional *blends.*

The two main approaches to presenting musical *expressivity* identified correspond directly to the *formal* and *referential* modes of listening; however, it has also been suggested that the most fruitful way to proceed may be to imagine emotional/feeling blends occurring across all modes. That is, the *expressive* mode of listening (housing emotions and feelings) might be best conceptualised as being *entangled* with all other modes of listening, producing the following two-way modal combinations or *blends:* expressive-formal, expressive-referential, expressive-visual, expressive-gestural, expressive-environmental, expressive-spatial, and expressive-structural. For the composer/sonic artist, this proposition presents much food for thought in terms of both established and novel ways to manage expression in the context of artmaking with sound.

Chapter Four: Formal Musical Qualities

This chapter is concerned with attributes of sound that have traditionally been regarded as *musically abstract*. Sonic characteristics of this type contribute to the formation of a functional musical *language* that can be interpreted by listeners through an extended range of recognisable patterns. Smalley (1996) refers to this mode of listening as the *interactive relationship.* For the purposes of this discussion, it has been re-termed the *formal mode of listening*. The areas of musical study encircled by this mode include vast subjects such as the rise and fall of Western harmony, and the theoretical underpinning of world rhythms. For practical reasons, only a brief overview of formal musical qualities will be offered with reference to melody, harmony, timbre, dynamics, texture, rhythm, and form - sound qualities and patterns that may be represented on the *pitch-time lattice*. These elementary sketches will then be cross-referenced to a defined set of *image schemas* provided by psychologists Lakoff and Johnson (1980, 1999), and Snyder (2000). This narrow approach to the topic of formal musical qualities attempts to conceptualise the subject within the psychological framework of image schemas, rather than within those supplied by traditional music theory. The rationale for this line of reasoning is to focus attention on identifying common patterns (revealed by image schemas) that may be integrated with other modes of listening.

Background

A large vocabulary of music-specific and metaphorical terms is routinely used to describe formal musical qualities. Every term within this expansive glossary relates to the *pitch-time lattice* - a conceptual framework that helps in understanding the domain. Each term serves to highlight a specific quality of music represented by the x-y grid. For example, the terms *interval* and *tempo* relate explicitly to each axis. In notated music (a representation of the pitch-time lattice), formal, expressive, and gestural dimensions are well represented through *the note* and the addition of an extended range of *expression markings* (also known as performance instructions) in the form of symbols and text

covering aspects like *dynamics*, *tempo,* and *articulation*. Expression markings may also specify nuance-level qualities of *character* that are best described through the emotion itself (for example, *dolce* – sweetly - or *agitato* - agitated). However, there are numerous important sonic characteristics that are not well represented by this construct - including four of the eight previously discussed modes of listening: referential, visual, environmental, and spatial (with some exceptions). Musical tradition largely underpins the concept of *formal musical qualities*. As such, the general term *music* or *musical sounds* is often used to describe formal musical qualities. This culturally informed understanding of what music is - and what it isn't - has both positive and negative connotations. Positive, in that we can rely on hundreds of years of scholarly thought together with a huge body of Western art music repertoire to help us understand the *functionality* of formal musical qualities. However, from a negative perspective, much of this understanding has come from outdated historical viewpoints. In the past, various qualities of sound - now known to be highly communicative, were considered non-musical. They were therefore excluded from traditional music-making practices, limiting the breadth and tools of sonic creativity. Traditional terms are, nevertheless, *useful.* This is especially so, if they are employed as part of a wider multimodal listening strategy that embraces an appreciation of *all* sound types and qualities.

The musical lattice consists of a vertical axis (containing pitch) and a horizontal (pertaining to time and/or rhythm). Some of the terms used to describe formal musical qualities speak specifically to these axes, but the majority describe combinations of both pitch and time. There are also various terms that naturally gravitate towards specific timeframes from the vertical combination of frequencies (*pitch*) to the phraseology or regular grouping of notes (*rhythm*) to the duration of entire works or even groups of works (*structure*). For the purposes of this discussion, a small cross-section of common terms pertaining to pitch, rhythm and structure will now be supplied: note position, harmonics, timbre, intervallic structure, vertical harmony, dynamics, scales and modes, pitch series, melody, harmonic progression, key, duration, tempo, time

signature, metre, pulse and pace, ostinato, phraseology, contour, cadence, repetition, themes, sections, parts, movements, proportion, ratio, and thread. More expansive lists are available through reference texts such as the *Oxford Dictionary of Musical Terms* (containing more than 2500 terms). Patterns relevant to melody, harmony, timbre, dynamics, texture, rhythm, and form will now be highlighted, discussed, and summarised with the aid of the pitch-time lattice.

Melody and Harmony

In the context of Western art music, the pitch class of sounds contains an in-built hierarchical structure where some pitches within a scale are phenomenologically appreciated as more important than others. The tonic is globally felt as a state of rest or resolution, the dominant provides a natural point of tension, the leading note and the subdominant also require resolution. Such detail is acquired culturally, and it is fair to say that these leading qualities are nowhere near as strong now as they might have been in the past. In modern times, we have come to accept a greater variety of melodic and harmonic progressions as normal; however, the tendency to regard the tonic as *home* remains a true universal across cultures.

It is commonly thought that the harmonic series may be at least partially responsible for this feeling of home, rest, or security in returning to the tonic. This pattern is represented in the consonance of the tonic chord I (in its major form), as it contains the fundamental and the first five partials of the harmonic series (which are mathematical multiples of the fundamental). The next three partials form 7^{th} and 9^{th} chords (8ve, 5^{th}, 8ve, 3^{rd}, 5^{th}, 7^{th}, 8ve, 9^{th}), which are, again, in popular use in several musical contexts. These observations related to pitch are well documented, carrying with them strong associations with abstract spatial constructs: perceived as rising and/or falling in vertical and horizontal patterns, defying or obeying gravitational forces, *building up* vertically (to form commonly used chords), and creating consonance and dissonance. The level of entanglement between formal and spatial modes of listening

is so strong that we cannot imagine a *high sound* originating from anywhere else other than the upper regions of localised space.

La perle et l'oubli (2002) - Gilles Gobeil

An example of expectations developed and thwarted within the context of this *pitch* class of sounds is an excerpt from Gillies Gobeil's *La perle et l'oubli* for onde martenot and tape. (1999-2002). I experienced this work live at the University of Montreal in 2003. Fifteen minutes into the piece the pitched materials settle to form a series of highly expressive and harmonically progressive chords. To end the phrase, a sequence of dominant acting chords is heard before returning to the tonic, gently at first, but then with unexpected fervour and earnestness. My initial experience was one of cathartic bliss together with bodily sensations of tension that were released when returning gently to the tonic, then, quite unexpectedly, my ecstasy was disrupted by the obviousness of the overstated tonic (which vulgarly restated itself again by soaring into the upper frequency range). To my astonishment, a large percentage of the audience were experiencing something similar – evidenced by laughing and shaking of heads. We had experienced a musical *joke* at our own expense – delivered by a master composer within the unusual context of experiencing live sonic art. After the concert I rushed up to the composer and awkwardly expressed my admiration. He confirmed that the joke was 'of course' deliberate. Joke-telling in musical contexts is not new – it dates back to at least Hadyn (1732-1809) whose speciality was 'musical buffoonery'. Gobeil used a similar technique in thwarting the audience's expectations surrounding musical form – in this case the perfect cadence provided the vehicle. The composer *led us down the garden path,* then *thumbed his nose* at us. His practical joke was audacious, but was hugely impressive in its delivery and timing.

Western art music has predominantly operated through the mainstay of pitch, where musical communicability is delivered through the structures

of *melody* and *tonal harmony*. Well-known aspects of the musical language achieved through pitch include: the ever-present tension, expectation of resolution, unexpected shifts to darker or lighter keys, the journeying through remote key centres in sequence and the way movements of departure are coupled with expectations of returning home. While it is true that novelty may now be hard to find in the pitch domain, it is one that continues to hold power across the general musical paradigm.

Timbre

The term *timbre* has traditionally been used to refer to 'the character or quality of music or voice as distinct from its pitch and intensity' (Oxford English Dictionary 3rd Ed.). In the context of instrumental music, it most commonly addresses differences in *tone quality* (for example, the *scraping* or *glassy* qualities of a note played on the violin). In 1994 Denis Smalley proposed a re-definition of the term as follows:

> … we must invoke a more abstract concept of timbre which can include the source. My definition therefore becomes a general, sonic physiognomy whose spectro-morphological ensemble permits the attribution of an identity. (Smalley 1994: 35)

Smalley's intention was to align the concept of *timbre* with his *spectro-morphological* listening strategy – a paradigm primarily based on the *referential* qualities of sound (subsequently expanded to the nine indicative fields: gesture, utterance, behaviour, energy and motion, object/substance, environment, vision and space). While this shift in thinking was important to the development of electroacoustic music, it may have come at a price. *Timbre* had long-since been, and is still, considered an *abstract musical quality*. In sonic art, *source-broken* materials (as opposed to *source-bonded*) tend to present themselves as relatively abstract. That is, although they contain timbral information, it is generally not explicit enough to achieve object or event recognition. This presents a continuum of referential associations emerging from the sound materials, ranging from implicit to explicit. These timbral qualities also correspond to the five mental layers of auditory attention:

sensations, gist, locus, contexts, and domains (as discussed in chapter two). Entanglement between formal and referential modes of listening may therefore occur anywhere along the implicit-explicit continuum, inclusive of the abstract end, where qualities such as pitch and rhythm tend to dominate over referential associations.

Dynamics

The concept of dynamics is useful when applied to instrumental sounds, particularly when considering their *articulation* through the associated concept of the ASR envelope (attack, sustain, and release values). This traditional formal musical quality, often referred to as 'dynamic expression' links with the *expressive* mode of listening. In spectromorphological contexts, the concept may be extended to control the loudness of individual frequencies as well as overall volume.

Texture

In orchestral music, the *term texture* usually refers to a mix of three or more instruments and more specifically, how different melodic, harmonic and rhythmic layers might combine to produce an overall sound described as homophonic, heterophonic or polyphonic. It may also be used to describe *textural change* between monophonic, heterophonic and polyphonic qualities. This traditional formal musical quality naturally pairs with several of the identified modes of listening: firstly - offering a counterpoint to *gesture* in a gesture-to-texture continuum, secondly - with the *environmental* mode through textural soundscapes that are outside the performative timeframes of instrumentalists, and thirdly - with *structural* thinking (as discussed in further depth within corresponding chapters of the book).

Rhythm

It can be said that any sound occurring along a given timeline can be considered *rhythmical;* however, the term carries further connotations such as the concept of *iteration* or repeated pattern. Musical *accent* also plays a major role in the appreciation of rhythm (particularly in drum music). There are hundreds, if not thousands, of established rhythmic

patterns the world over. These patterns, both simple and complex, might also exhibit variations or developments along certain pathways and/or towards certain goals. If the intended goal is recognised by listeners before arrival, then composers/performers might in this way create musical expectations. Such pathways also serve the overall structures of sonic artworks.

Form

The concept of *form and structure* has many dimensions. On a large scale, it concerns the *structural integration* of works and the organisation of sonic materials. (This topic will be addressed in detail in *Chapter Ten: Structural Concepts.)* Structure can also supply *meaning* as it lends itself well to dramaturgy and quasi-narrative – both of which play active roles in the assignment of referential meaning. For the time being, we will focus on the formation of musical patterns and the creation of *component structures* within the narrow context of working with formal musical qualities. This topic may be conceptualised by way of the interplay of three related structural principles proposed by John Young (2016): moment form, morphic form, and narrative.

Moment form refers to the psychological impression of objects or events that make themselves apparent during the ordinary listening process. They may be conceptualised as having spatial and temporal boundaries and play a part in forming the structures of music through the way in which they are attended and remembered. Stockhausen defines the term as follows:

> When certain characteristics remain constant for a while – in musical terms, when sounds occupy a particular region, a certain register, or stay within a particular dynamic or maintain a certain average speed- then a moment is going on: these constant characteristics determine the moment (Young 2016: 64) referencing Stockhausen 1991:63)

Moment form arguably requires *precision timing* to function to its full capacity. Nuance-level details such as ASR (attack, sustain, release) curves, perceived agential energies, as well as periods of proportional

silence to follow, tend to combine to produce the most convincing results. Moment form was originally conceived within the context of traditional musical materials; however, some leading researchers claim that the concept also extends to sound-based materials – as John Young attests to:

> In continuing this line of thought, I am going to suggest here that a recognized *source* may have a moment function, not solely in spectro-morphological terms, but also in referential terms. Young 2016: 69

Morphic form concerns impressions of progressive change that develop as the music unfolds. The concept of *continuum* is essential in defining its boundaries - for example, soft to loud dynamics, high to low pitch, or thin to thick timbre. While these qualities are important to the structural development of all musical genres and categories, acousmatic music features a genre-specific variety of morphic form that is worthy of special mention - the technique of spectral metamorphosis commonly referred to as 'morphing' - as John Young explains:

> Trevor Wishart's (1994a) genre of 'morphic' form extends from the potential of electroacoustic music to develop materials through processes of continuous transformation of sound identity and timbre. Sound transformation tools found in electroacoustic music may encourage this kind of approach since notions of sonic transfiguration and metamorphosis can take on new, causally centred meaning (in the sense of sound as aural product of interactions between a sounding body and some energy input). (Young 2016: 67)

Reference to 'causality' indicates that laws relating to the physicality of the natural world have also made their way into this highly generalised structural concept. One such rule is the law of conservation of momentum. While it is easy to imagine the stroboscopic motion of a series of abstract musical pitches, it is more difficult to imagine continuity and flow within the context of sound-based music. As Gary Kendall explains:

> While a composition is unfolding in time, the listener is always making connections, forming and releasing mental spaces, and attempting to make sense of the unfolding experience. But what guides the listener to establish a thread of continuity like the metaphorical journey, linking the mental space of one event to another? Why and how does the listener string together particular events from among many simultaneous and overlapping events? In everyday life, we experience continuity among events that have particular things in common, usually things that have common consequences. For example, while driving in the car, everything having to do with traffic is one thread of continuity and the conversation on the radio is another. Listeners to electroacoustic music could establish continuity in similar ways. … Source, flow dynamics, agency, place – these are all factors that might influence the listener's lines of continuity. (Kendall 2010a: 72)

The final broad category concerns *Narrative.* The term embraces all forms of referential meaning including indicative qualities (Smalley 1996), musical dramaturgy, and storytelling presented through a variety of materials including semantic sounds (such as spoken voice). Interestingly, the concept of narrative appears to assist in making sense of structure even in the most abstract of musical settings. Again, John Young provides the following useful summary statement:

> Conceptions of narrative are central to time-based arts. Barthes (1977) regarded narrative as an 'international, transhistorical, transcultural phenomenon, simply there, like life itself': we find narratives because we view ourselves situated in a present, with a past and designs for a future and it is difficult to avoid a conviction that it is a fundamental dimension of human thought and understanding (Young 2016: 72)

Viewed through the lens of the *structural mode of listening,* these related concepts are observed either in isolation or combination as music progresses through time. In essence, they are understood as abstracted (drawn from the materials) or abstract (drawn from elsewhere) patterns that access a variety of structurally associated schemas stored in the short and long-term memories of listeners.

Summary: The Musical Lattice

This brief overview of *pitch*, *rhythm* and *structure* is not intended to be comprehensive. It, merely, provides a succinct mental impression of musical patterns as they might be represented on the pitch/time *lattice*. It is also important to iterate that these qualities may be attended in *all* sounds, not just abstract instrumental and vocal materials. In summary (as a conceptual overview), low-to-high pitches are assigned to the vertical axis (melodies are represented as sweeping contours and chords organised in vertical bands) and their horizonal positions may be considered *rhythmical* along with other non-pitched iterative and accent-based percussive materials. *Structural* timeframes are highlighted by sounds that are particularly striking or memorable, the repetition of patterns, and/or developmental changes. It is this type of *spatial* appreciation of musical patterns that lends itself well to cross referencing with *image schemas.*

Musical Patterns as Image Schemas

In their book *Philosophy in the Flesh* (1999) George Lakoff and Mark Johnson present a description of *image schemas* based on their conviction that there are only 'several common types' of archetypal guiding principles formed in the human brain through processes of spatial reduction. According to Johnson's 1987 book, *The Body in the Mind,* an image schema is:

> a recuring, dynamic pattern of our perceptual interactions and motor programmes that gives coherence and structure to our experience. (Johnson 1987: xiv)

Such abstract mental structures (as few as five main types) are expected to bind with early sensory data (largely visual and auditory) to create broad categories of what to expect from the experience of attending any given sound, object, or event. One can imagine that such quick responses may have formed as part of our defence mechanisms in the earlier stages of evolution. Although some theorists criticise the grandeur of this suggestion due to the absence of any supporting evidence, Lakoff and Johnson propose that there *is* some evidence to support the claim in the

form of *psychophysical complementarity* (shared spatial reductions relating to image schemas, various qualities of the human body, and physical constants of the universe). The bodily sensation of gravity, for example, might be directly paired with the up/down image schema. They state:

> We experience space as structured by image schemas (as having bounded regions, paths, centers and peripheries, objects with fronts and backs, regions above, below, and beside things). Yet we now know that space in itself has no such structure. The topographic maps of the visual field, the orientation-sensitive cells, and other highly structured neural systems in our brains not only create image-schematic concepts for us but also create the experience of space as structured according to those image schemas. (Lakoff and Johnson 1999: 452)

Such a description allows an appreciation of image schemas in terms of both *patterns* (abstract constructions of the world open to cognitive scrutiny) and *feelings* (as experienced in the human body). Examples of patterns within the context of music include *pitched* sounds that are perceived to be *higher* or *lower* than others; rhythmic or scalic patterns that are *inside* or *outside* their bounds (for example: *in time,* or *out of key*), while *structurally,* they may be perceived to be part of a development towards a particular goal. These patterns also have associated *feelings* centred in the body (for example: sensations of *tension and release*, gratifying feelings of being *inside* the music through precise rhythmic synchronicity, and sentiments of *longing* or *striving* to reach specific goals). By reducing these spatial categories to a minimum, we are left with only a few conceptual frameworks that can contain all musical patterns. It is claimed that such constructs are so broad in their scope that only a small collection of different types of image schemas can, in combination, offer a complete spatial reduction of all human experience. Lakoff and Johnson offer some graphic figures to assist in the process of conceptualising the various types of image schemas proposed. While there is some resistance (on their part) to defining the total number and specific types of image schemas in absolute terms, there are

some categories that are plainly identified. Two examples are offered below (Lakoff and Johnson 1999: 39-40):

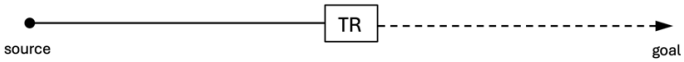


Fig. 4.1: Source-Path-Goal Schema

The source-path-goal image schema refers to the experience of making accurate predictions about the future based on recuring or continuous patterns. The TR in figure 4.1 refers to the term *trajectory* – a metaphor that may be applied to a developmental pathway. The metaphor may also be used as an example - a ballistic missile, once fired, is likely to reach its goal.

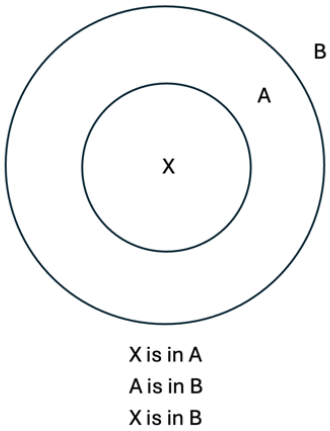


Fig. 4.2: Container Schema Logic

The container image schema refers to understandings of certain things being *inside or outside* other things. X, A, and B represent any given variable that suits the construct. For example, John (X) is in his house (A). John's house (A) is in New Zealand (B). Therefore (logically), John (X) is in New Zealand (B).

Bob Snyder wrote extensively on the topic of image schemas in *Music and Memory* (2000). The significance of his work cannot be overstated, as he identifies and specifies a series of five guiding principles (presented below) that operate in the service of musical communicability, with a specific focus on *formal musical qualities*. He claims that these *image schemas* are the scaffolding on which both spoken and musical languages are built. He writes:

> Recent theory has suggested that metaphorical mappings are not arbitrary but are grounded in fundamental embodied cognitive structures generalised from recurring physical experiences, especially the experience of our own bodies. These cognitive structures have been referred to as image schemas. Image schemas are thought to be derived from commonalities in different experiences that seem related, such as, they are believed to form a basis for our conceptual systems, indeed, to connect our perceptual experience and concepts. Image schemas represent the most stable constancies and structures we all share as human beings, derived from dynamic patterns of interaction with our environment…The actual cognitive form that image schemas are thought to take is somewhere between concrete, specific visual images and abstract concepts. That is, image schemas are often derived from types of situations that we can visually imagine, but that are more abstract than any particular image. …. This is another way in which image schemas are different from either visual images or abstract concepts- they can have a kinaesthetic component and represent muscular sensations in relation to particular experiences; they can have a particular physical "feel" to them (Snyder 2000: 108).

In providing this background information, Snyder intentionally steers the research paradigm of *image schemas* towards the context of appreciating music. In doing so, he is proposing that the experience of listening to music is *multisensory,* involving at least the construction of quasi-visual and spatial *images*. He also supports Lakoff and Johnson's claims that abstract patterns (based on spatial constructs) are normally accompanied by sensations within the human body that might be defined by physical *feel.* Gary Kendall (2010) also supports this by directly addressing the proposed relationship between feelings and the human body (initially supplied by Johnson) within the context of listening to music:

> In everyday life, the experience of events is interwoven with the flow of felt experience. Similarly, Johnson, speaking of the embodied nature of image schemas, explains that bodily movements are characterised not only by their image schemas, but also by their 'distinctive qualities' (Johnson 2007: 21). These he relates to the four dimensions of body movement described by Sheets-Johnstone (1999): tension, linearity, amplitude and projection. This terminology could equally well be applied to the flow of energy in musical events. (Kendall 2010a: 66)

Image Schemas and Formal Musical Qualities

The image schemas referred to in this section (proposed by Snyder 2000) may also be considered as metaphors to aid in the understanding of formal musical qualities and how they function. Five categories are identified: up and down, centrality, motion-linkage-causation, linearity: pathways and goals, and containment: inside and outside. It is submitted that understanding possible metaphorical connections between music and experience can help us not only to understand music, but also, create it. The following presents a means of categorising the Snyder's observations in relation to the five image schemas identified. (I have taken the liberty of paraphrasing, and freely elaborating on his comments sourced from Snyder 2000, pages 111-116).

Up and Down

'In general, increases in energy are thought of as *rises* and decreases as *falls'*, and with this in mind, moving down is commonly appreciated as easier than moving up. Furthermore, 'descending gestures of closure in melodic patterns (lowering of pitch) may be extended to descending gestures of closure in other parameters, such as loudness or timbre to serve different aesthetic needs. Intensity is broadly defined as any change in a stimulus that causes an *increase* in neural activity. By this definition, any change in sound - not just loudness - that is perceived as an *increase* makes the sound more intense. 'This includes making the sound higher in pitch, harmonically more dissonant, brighter in tone colour, faster in speed, or denser in temporal distribution'. The up/down image schema is inherently linked to the concept of gravity - a body-centred sensation

that may endorse observed rules such as *all that goes up must eventually come down*).

Centrality

'The closer together two patterns appear in a piece of music, the easier it is to recognise their similarity, hence even though they have differences, they are still recognised as similar. On the other hand, two patterns that are far apart in a piece of music must be very similar for their similarity to be apparent.' This concept is commonly observed within the many structural forms of music (for example, theme and variations, sonata form, verse-chorus form, etc.). Furthermore, 'repeatedly returning to a particular pitch or pattern creates a closural effect. This is especially true of returning to the central pitch of a tonality, a form of closure that involves the memory of other occurrences of that pitch. This memory does not appear to have been consciously learned and is an example of implicit memory.' In Western art music, the phenomenon of pitch *centre* has developed over the centuries into a sophisticated tonal language incorporating, not just a hierarchical intervallic pitch series, but also associated and remote *key centres*. When we transfer the idea of centrality to the concept of *space*, we also create the possibility of moving away from and toward a central location. 'Physical centrality is based on the ideas that some places are somehow more important than others, and that is why we return to these places again and again. These kinds of places are often marked or emphasised in some way. Musical centrality is based on these ideas as well. Because music is a temporal medium, musical centrality is often developed through repetition. Metaphors of musical centrality can refer to *central pitches, central beats* in metrical structure, and/or *central landmark events* in a piece, all of which gain their centrality by being repeated at different times.'

Motion-Linkage-Causation

The principle of continuity (drawn from Newtonian law) states that consistent movement in a particular direction tends to perpetuate itself in the same direction. In music, tones or rhythms are said to be *in motion*

– 'this is somewhat like visual stroboscopic *motion* perceived when separate lights near each other are turned on and off in sequence and appear to be one light that is moving.' When it comes to pitch, small intervals tend to create the strongest sense of motion. Interestingly, 'repetition is not progressive and hence does not constitute musical motion'. This observation (based on several psychological tests) may help to explain the functionality of certain musical genres such as *minimalism,* where small changes within a greater morass of repeated patterns are often held responsible for the sensation of forward motion in the music. Finally, 'the metaphor of musical motion implies a complimentary metaphor - the lack of musical motion, or *rest*.' Lakoff and Johnson point out that 'listeners have an understanding of spatial motion that is deeply connected with sensorimotor experience.' (Lakoff and Johnson 1999: 179) Spatial motion follows paths, often encounters resistance and requires some form of energy. Such understandings are *metaphorical,* which brings to light the quasi-narrative (storytelling) potential of image schemas.

Linearity: Paths and Goals

This image schema embodies the principle of expectations formed around the structures of musical works. Such impressions are primarily *metaphorical* arriving through attempting to appreciate one thing in terms of the known patterns of another. 'Arguments, for example, are often metaphorically thought of as journeys that have a *line*, parts that *follow* each other, from which we get *side-tracked* and so on. We also use metaphors such as *getting somewhere* and *things not being there yet* to structure many types of conceptual *problem spaces* as though they involved passage through physical space toward a goal. Like the other image schemas, the pursuit of goals is an experience so fundamental it is probably shared by human beings of many times and cultures. The idea of a goal implies that the present is somehow structured by an idea of how the future will or should be.' In terms of forming expectations in the minds of listeners, 'a pattern of continuous change of musical parameters in equal increments is not very interesting. Indeed, even though most interesting music sets up patterns of expectation using at least some

consistent change, it deviates enough from this pattern (putting the goal in question) to reduce predictability and maintain attention.' As a general statement, 'the more parameters that reach closure at a given point, the higher the structural level of closure achieved at that point.' This final comment may help in conceptualising goals related to cadencing or ending creative works, and in ensuring that all materials have been reconciled.

Containment: Inside and Outside

There are many examples of containers that are in popular use within the context of English language. For example, in love, out of touch, on the edge, out of it, etc. From a musical perspective, 'saying that a particular note is *in* a particular phrase, for example, implies that the phrase has boundaries at its beginning and end points.' Other common examples relating to pitch, rhythm and structure include concepts of being *in* or *out of* tune, range, intonation, key, time, sync, sequence, style, genre and more generally *in* the zone, or *out of* practice. 'Another metaphorical type of boundary is involved in the concepts of surface and depth. The surface (or just below) delineates the boundary between.' *Good music* is often identified as having both *surface* and *deep-level* features.

To summarise this section, the overlay of formal musical qualities with image schemas has highlighted a range of *formal artmaking strategies* that will be examined further in the ensuing section. It has also called to attention the co-existence of numerous modes of listening acting in *combination* with formal musical qualities that will be subsequently discussed.

Formal Artmaking Strategies

Following Snyder's (2000) lead, in overlaying formal musical patterns with the five image schemas identified: up and down, centrality, motion-linkage-causation, linearity: pathways and goals, and containment: inside and outside, a surprising range of *formal artmaking strategies* comes to light. These strategies, arriving through *observation* may also be used intentionally for compositional, developmental and/or structural

purposes. As the patterns have been identified as *archetypal* (possibly related to the principle of psychophysical complementarity), there appears to be very little risk of exposing artificiality or inauthenticity through their use. Specific artmaking strategies include:

- *rises and falls* in any sonic parameter to emphasise an increase or decrease in 'neural activity.'
- the principle of *similarity* applied to sound objects, events, and tonal centres to provide a sensation of *returning home* with the principle of *proximity* also playing a role in the appreciation of similar sounds, patterns, keys, etc.
- the sensation of *movement* that might be elicited through *stepwise melody* as well as the *metaphorical associations* with a variety of movement types.
- the *trajectories* and *continuums* and *expectations* associated with structural *pathways and goals.*
- the *boundaries* (inside and out) of tune, range, intonation, key, time, sync, sequence, style, and genre.

These are but a few of the potentials exposed. Highlighting the subtopic of formal structural qualities, the above examples also evidence an overt link between image schemas, structures and the human body. As Lakoff and Johnson explain:

> Conceptual structure arises from our sensorimotor experience and the neural structures that give rise to it. The very notion of "structure" in our conceptual system is characterized by such things as image schemas and motor schemas. Mental structures are intrinsically meaningful by virtue of their connection to our bodies and our embodied experience. They cannot be characterized adequately by meaningless symbols. (Lakoff and Johnson 1999: 77)

Concluding Remarks

The overlay of formal musical qualities with image schemas has highlighted naturally emergent multimodal relationships (between formal and other modes of listening). Through an investigation of the communicative potential of image schemas, it has been found that they induce in listeners, at least bodily sensations and metaphorical associations. *Bodily sensations* (for example, muscular tension and release) might be further understood as emotions and/or feelings housed by the *expressive* mode of listening, as well as gravitational sensations highlighted by the *gestural* mode of listening. This description presents both *formal-expressive* and *formal-gestural* modal combinations that may be seen to be relevant to a large percentage of Western classical music. Formal-gestural blends also emphasise the role of the human body in expressing musical phraseology, particularly when considering the attack-sustain-release (ASR) curves of articulation and their agential relationship to the performer. Metaphorical understandings of music (for example, associations with known events such as *journeying*) highlight other common modal *blends* such as *formal-referential,* and *formal-structural* combinations. The natural association of images schemas with visual and spatial constructs also works well to highlight the remaining three modes of listening. *Formal-visual* combinations highlight situations where formal musical qualities might be appreciated through the quasi-visual images they evoke (for example, melody *lines*, or structural *pathways*). *Formal-environmental* combinations call to attention a range of high-level universal laws (such as the conservation of energy and momentum) that still apply even in the most abstract of musical settings. *Formal-spatial* combinations accentuate a variety of abstract spatial forms complete with dimensions and territories (such as high/low sounds, or expectations formed through developmental pathways).

Worthy of special mention is the domain of sound-based music and acousmatic music (particularly), which offers enormous potential to composers and sonic artists through the vast domain of *formal-referential* blends (formal and referential modes of listening in combination). The relationship supplies the 'dual potential' that Smalley

referred to in 1986 and viewed through this lens, the *zeitgeist* of much electroacoustic music becomes apparent. This specific combination of modes was/is *so inviting*, that it gave rise to an entire musical paradigm where the musical qualities of *all sounds* (not just instrumental sounds) might be perused and enjoyed.

Given the expressive, referential, visual, gestural, environmental, spatial, and structural associations that formal musical qualities are able to present (through the eight entangled modes of listening identified) together with the fact that they may be perused and enjoyed in their own rite (solely through the formal listening mode), it is hardly surprising that sophisticated communication of human ideas has been experienced through abstract music - across oceans, cultures, and periods throughout history. More than 200 years ago, Henry Wadsworth Longfellow wrote, 'music is the universal language of mankind.' At that time this belief was so strong, that music was often used as a means of establishing *first contact* with peoples of diverse races and cultures the world over. While this approach is now known to be overreaching, it is also known that formal musical qualities have wide communicative potential – especially when combined with other modes of listening.

Chapter Five: Referential Meaning in Music

Meaning in music is a concept that is difficult to pin down. It has many contributing dimensions, is highly contextual, and tends to be attained through a multitude of subjective interpretations. This indeterminate hallmark of music is of course one of its very special qualities that contributes to its status as an *art form;* however, its non-specific nature can be also challenging to manage within the compositional process. On one hand, musical meaning can provide profound insight, while on the other, it can mean entirely different things to different people. Gary Kendall provides an excellent summary of this somewhat paradoxical issue in the following comment:

> Meaning is a product of the individual listener's mental processes: whatever meaning the listener makes of the listening experience, that is its meaning to the listener… Therefore, the analyst should be mindful not to ascribe meaning to the physical acoustic attributes of a work or to graphic representations based on signal analysis in which the listening experience is not captured. For analysis to be truly credible, it must be envisioned fundamentally in terms of human perception and cognition. This places a rigorous responsibility on our analytical techniques to justify their relevance to the listener's mental activity and to avoid unexamined implicit assumptions about what is meaningful to the listener. (Kendall 2016: 31-32)

Kendall's statement highlights the *contextual* nature of meaning as well as what he sees as a necessity to cross reference any claims concerning the assignment of meaning with known scientific understandings of the process of human audition. Although it is not directly stated, the comment also challenges the common approach to meaning adopted by the international electroacoustic music community, which is primarily based on Schaeffer's mode one (Écouter), and Smalley's indicative relationship. While this message-bearing mode is important, and offers much to the current discussion, it forms only a component of the total musical meaning. The wide approach adopted in this chapter is therefore to consider *meaning* as a product of all identified modes of listening

(expressive, formal, referential, visual, gestural, environmental, spatial, and structural) supported by relevant research from the field of psychology (such as the five layers of mental activity model already discussed). This broad definition of *meaning,* embracing all known modes of auditory attention, logically develops through reception of all *information* delivered through sound – a paradigm that has already been examined at depth in *Chapter Two: Multimodal Listening.*

While this wide interpretation may be useful in terms of emphasising the wide range of meanings that can develop from the listening experience, the more common understanding of meaning in music, tends to connote limiting terms such as *expressive,* or *referential.* In *Chapter Three: Music as Expression*, these two modes of listening (expressive and referential) were briefly discussed in terms of the *meaning* they convey in music. The latter of these two subtopics (*referential meaning*) will now be examined at depth. It constitutes the narrow focus of the chapter, circumscribing the sonic characteristics that might encourage object or event recognition. These abstractions include Smalley's (1996) nine indicative fields, semantic qualities, signs and symbols, images schemas, and the principle of conceptual blending.

The chapter is therefore loosely divided into two parts. The first part clarifies the way in which contributing factors operate in the assignment of *referential meaning* (distinct from *expressive* or other forms of meaning), while the wider and more general topic of *meaning through entanglement* of the eight modes of listening identified is primarily addressed in the latter half of the chapter. The second part of the chapter also presents a small range of practical *meaning-based artmaking strategies.* Three interrelated techniques are proposed. These include augmentation of meaning-based qualities, justification of techniques, and conceptual blending as meaning.

Referential Sound Qualities as Meaning

The referential mode of listening differs from others discussed so far in that literal or referential meaning is often far less open to interpretation.

It is formed through associations with objects or events from life experiences that may be shared across groups and cultures - but may also involve the decoding of semantic materials (for example, a spoken story), and in these settings, those who understand the association or language often arrive at common or even singular understandings of meaning. These *referential sound qualities* may be succinctly defined as *having the quality of a 'sign'* (a term borrowed from the field of semiotics) - presenting the view that referential and semantic materials are able to act as signifiers of something external to the music. The way in which *referential meaning* is assigned is primarily due to the phenomenon that: some sounds are perceived as being more important than others in terms of the associative information they present. Developing an understanding of these principles provides composers and sonic artists with a set of tools for managing the assignment of referential meaning in their works. This is achieved through the deliberate and/or intuitive use of what Smalley (1996) refers to as the 'indicative qualities' of musical sounds - relying on logic, reason, and deduction – with reference to external knowledge coupled with the creative imagination of both the composer/performer and listener. Smalley claims that referential associations are formed through an ensemble of *indicative fields* that extend throughout the concrete-abstract continuum. It is suggested that *discovery* of these qualities together with the listening *imagination* play major roles in the assignment of referential meaning. Just the right amount of information is required to allow the composer's message, or something close to it, to be apprehended during the ordinary listening process. Smalley breaks down this sonic information into nine contributing fields. He states:

> Nine fields are identified. Three are archetypal: gesture, utterance and behaviour. These fields are original universals. Human utterance and the consequences of gesture have traditionally provided the fundamental sounding models for music. The behaviour field is concerned with sounding relationships in space and time, which can be considered analogous to certain modes of human relationship, observed relationships among things or objects, or human-object relations. The six remaining fields are energy, motion, object/ substance, environment, vision and space. A single field is not autonomous and does not exist either inside or outside music without

> overlapping other fields. The notion of the indicative field/network comprehends relationships that are considered mimetic. (Smalley 1996: 83)

The following paragraphs present the numerous observations made by Smalley in relation to the nine indicative fields discussed. I have taken the liberty of paraphrasing and freely elaborating on his comments primarily sourced from his article *The Listening Imagination: Listening in the Electroacoustic Era in* Contemporary Music Review Vol 13 Part 2 1996 (pages 77-107). Those that are used verbatim are presented in quotation marks.

Gesture/Texture

The gesture field houses psychological associations that the music/sound is created by human physical activity. Gestures are also responsible for the assignment of identifiable agents (real or imagined) ranging from clearly identifiable human sources (first-order surrogacy) to more elliptical associations (second-order surrogacy) to remote surrogacy, where 'the indicative link can be forged through the energy-motion trajectory alone, without reference to real or surmised physical gesture or an identifiable source'. Gestures operate within very specific timeframes, which may be further understood through the overlay of attack, sustain, release (ASR) envelopes that correspond to human physical movements (such as the movement of limbs or breathing). Beyond these timeframes (a maximum of say 15 seconds) is the domain of *texture. Textural* appreciation involves a shift in auditory interest from the ASR envelope of the sound (perceived as a single sonic event) to 'internal spectral motion, which may either reinforce or contradict the external gestural motion'. Smalley has been criticised by some theorists for his omission of non-human agents that might be perceived as identifiable sources (for example, mechanical or electrical) but perhaps this is part of the point he is trying to make – when it comes to second order, and particularly remote surrogacy, it is our tendency to *anthropomorphise* sounds, and in so doing we routinely create associations with imagined performative agents. He states, 'the listener is thus called upon to exercise and enjoy maximum gestural imagination'.

Utterance

The utterance field houses all associations with the human voice (real or imagined) 'The fact that the sounds of utterance are generated from within the body and that they are the essential vehicle of personal expression and communication, makes utterance intimate and emotionally charged. Therefore, the listener's relationship with utterance is often reflexive rather than indicative'. This is one of the fields in which the true level of entanglement between different types of meaning becomes apparent – although we are currently discussing referential qualities with the human voice, common reactions to these materials tend to be emotionally-centred and may be better represented through the *expressive* mode of listening (particularly when it comes to more abstract vocal gestures). The presence of *utterance* materials also changes the context and functionality of surrounding materials in that the listener's attention is often drawn to 'indicative meanings associated with the perceived utterance'. The wide scope of the category embraces both nuance-level changes in pitch, volume, timbre, and rhythm, and structural-level narrative devices drawn from the domain of storytelling.

Behaviour

'The presentation of two or more sounds together automatically ensures that relationships will develop between them. The term 'behaviour' has been chosen to represent these relationships.' Three types have been identified: dominance-subordination, conflict-coexistence, and causality. 'Dominance-subordination is concerned with the (in)equality of participating morphologies and is related to three interlocked ways of interpreting the idea of foreground and background: spectro-morphologically, spatially and indicatively. A simple, obvious example would be bright, spatially present, impact-attack morphologies superimposed on a sustained, relative distant, spectrally inactive morphology'. The indicative impact of a sound can also affects its dominance: 'we are already aware of the voice's focal pre-eminence'. Identified sources also tend to dominate the unrecognised. 'Conflict/coexistence is intended to highlight temporal relationships within the behaviour field. On the one hand sounds can exhibit

competitive tendencies, and on the other they can exist in a reciprocal, confluent environment. The third major area of behaviour is *causality* which, although it exhibits attributes of the two paired oppositions, needs separate consideration. Causality in this context does not refer narrowly to physical gesture. It is more concerned with one sound's action upon another, either causing the second event to occur or instigating change in an ongoing sound.'

Energy and Motion

'The energy and motion fields are intertwined in the spectromorphology of sounds.' They are *abstract* in nature and bear obvious relationships with the five image schemas already discussed: up and down, centrality, motion- linkage-causation, linearity: pathways and goals, and containment: inside and outside. This is yet another field displaying multiple levels of entanglement – with at least *formal, referential*, and *spatial* modes of listening. 'Musical motion will be interpreted as moving forward if a directional tendency is sufficiently strong and/or is maintained for an adequately long timespan. Directionality is related to goal orientation and growth as in converging/diverging, ascending/descending contours, and accumulative/dissipative and diffractive/conglomerative growth'. The category also embraces the concept of expectations fulfilled and thwarted through the appreciation of musical structure as energy and motion pathways.

Object/Substance

'The word 'object' is commonly used in the term 'sound object' [a Schaefferian concept] although it is used somewhat loosely'. Smalley intended to broaden this concept through the accompanying term *substance* for the purpose of embracing 'thingness'. The category highlights material-like qualities within the spectromorphology of sounds that play roles in defining what the object, substance or event might be. The field defines the *what-like* qualities of sounds, whilst recognising the fluidity of such associations within the context of both listening to electroacoustic music (where these qualities might be deliberately manipulated), and as part of the general listening process.

Environment

In chapter two, the fundamental nature of relationship between humans and the natural environment was highlighted as one that ultimately supplies the framework on which the listening experience operates. It has therefore been deduced that environmental associations may be considered as being active throughout *all* modes of listening. Within the current context, it is Smalley's intention to highlight that 'the incorporation of environmental sounds into music via the recording medium has led to a greater appreciation of the musical qualities of sounds outside the immediate human orbit.' His inclusion of *the environment* appears to be a literal one, where soundscapes of all descriptions might be superimposed on note and/or sound-based musical textures not only to provide a background but also to highlight musical qualities within the environmental recordings themselves. The environment field also provides a natural counterpoint to the *gesture* field, and in this way may be seen to operate in parallel to Smalley's concept of *texture.* That is, while gestural activity may be found within environmental contexts, environmental sounds are often perceived as *textural* (operating in background spaces).

Vision

'Every field discussed so far contributes to the visual indicative network. It is true to say that vision is at the very basis of the gesture-field, and that the energy-motion trajectory is unimaginable without its visual correlations. This would imply that music, and electroacoustic music in particular, is not a purely auditory art but a more integrated, audiovisual art, albeit that the visual aspect is frequently invisible. That in turn suggests some kind of synaesthesia'. The insight Smalley offers in this succinct statement is extremely important, given that we are only now, thirty years later, beginning to come to terms with the true multimodal and multisensory nature of the listening experience.

Space

In his 1996 article, Smalley begins to break down the space field into distinct categories by first stipulating 'an enhancement or reinforcement

of sounding properties inherent in spectromorphologies and structural relations'. Within this category, rather than remain fixed on his description of spectromorphology as a listening strategy, he instead adopts a poietic (composer's) stance that appears to connote the ideal presentation of acousmatic work through multichannel loudspeaker arrays. Next, he stipulates an area concerning 'sound diffusion - the articulation of the composed spatial content in public performance' (relating to the concert hall), and thirdly an area that 'concerns the affective interpretation of space, and how the listener experiences and feels about space.' Daniel Barriero provides a useful summary of the essential qualities of this final category (Space) as 'related to the perception of intimacy and immensity, confinement and vastness, among other references related to human experience.' (Barriero 2010: 37). Smalley also discusses three conceptual categories of 'relational structure. The first is the relationships and behaviour among sounds within the composed musical space; the second is the movement between successive spaces, or the transformation of space in the work; the third is the interaction between the musical space and the listening space that for the listener will be fused with the first two.' Smalley expands on these comments significantly in his 2007 article *Space Form and the Acousmatic Image* discussed in detail in *Chapter Nine: Spatial Forms.*

In summary, the nine indicative fields impact on the listening process through the component assignment of referential meaning(s). They present the literal material-like *what* characteristics, in combination with the feelings, formal qualities, images, gestures, environments, spaces, and structural forms associated with the objects/events they represent. They may do this in overt or subtle ways, not always source-bonded (another term proposed by Smalley), but sometimes offering associations with multiple possible sources that might result in what Smalley refers to as *bonding play.* The sound of gravel falling might, for example, evoke associations with falling rain, running water, ocean waves, or avalanches. Any one of these associations might be considered as a *correct* interpretation. The term *affordance* is useful in further explaining the

connection between literal (*what*) associations and meanings. According to Luke Windsor, the word *affordance*, originally proposed by Gibson (1979) may be applied where 'the dynamic relationship between a perceiving, acting organism and its environment is seen to provide the grounds for the direct perception of meaning' (Windsor 1995: 68). Gibson Provides the following definition:

> Objects and events are related to a perceiving organism by structured information, and they "afford" certain possibilities for action relative to an organism. For example, a cup affords drinking, the ground, walking. For different organisms, affordances will differ; for a human being an open body of water might afford swimming or immersion, but for a water bug the same stimulus information, picked up by different perceptual systems and relative to different organismic structure, would afford support (Gibson, 1979, p. 127).

When applied within the context of listening to electroacoustic music and more specifically when attending Smalley's nine indicative qualities of sounds, the process of recognition of certain objects or events *affords* the listener certain possibilities. The recognition of ocean surf, for example, might *afford* some listeners associations with summer holidays, as well as feelings of relaxation, while, for other listeners who recognised the same sound as [say] an avalanche, the *affordances,* and therefore the associated meaning(s), might be entirely different.

Semantic Qualities

Still within the boundaries of Smalley's nine fields – and specifically within the *utterance* field, we find spoken languages. These sounds work in ways that we are all familiar with. They are encoded, through speech, to mean specific things. They are *semantic* – indicating that they specifically relate to the concept of meaning in *spoken language.* Communication relies on auditory attention being primarily focussed on these encoded associative qualities. The word *banana*, for example conjures up images of the fruit itself, but also recalls more distant concepts related to bananas. In musical contexts, words may also be used to present narratives or poetry – as is the case in much folk and popular music. The domain of

programmatic sonic art too, offers much potential for storytelling. The *Electroacoustic Resources Site* (EARS) website states: 'electroacoustic music lends itself particularly well to semiotic study where there is a presence of identifiable sounds.' (EARS accessed 18:02:2025). Here, the use of literary devices such as imagery, symbolism, flashbacks, foreshadowing, allegory, juxtaposition, etc. are made available to the composer in the telling of complex and emotive stories drawn from human life experiences. The use of semantic materials not only opens the way to the use of storytelling techniques but also offers up new potentials to *connect* with audiences and individuals *through* the story itself that might resonate with aspects of the listener's personal life experience. Within the context of acousmatic music this most often relies on the use of prerecorded materials - as John Young states: 'The view of narrative presented here starts with the capacity for sound recording to function as a mirror held up to lived experience.' (Young 2016: 72).

Signs and Symbols as Meaning

The assignation of meaning can be understood further by utilising semiotics theory to understand its constituent parts and how they accumulate. It also helps explain how individuals might gain divergent meanings from the same materials and how composers may not be fully aware of the depth of possibilities that listeners might gain from a work. Semiotics theory uses the edifices of *sign, signifier* and *signified* to illuminate the referential meaning-making process for both composers and listeners when presenting/receiving a work of art. The Semiotic Triangle (presented in figure 5.1) explains the terms *sign, signifier* and *signified*. The apple object pictured in the model is *the sign* (object or thing), a term that may be interpreted literally as a *signpost* pointing to the concept or *meaning* of the apple (*the signified)*, as fruit, healthy food, and everything else associated with apples: temptation, poison apples, the big apple and Apple computers. Returning to a previously used term, the sign *affords* these possibilities to the listener. *The signifiers*: (for example, the word 'apple') are in physical existence but are considered as a collection of parameters independent from yet pointing to the object of perception.

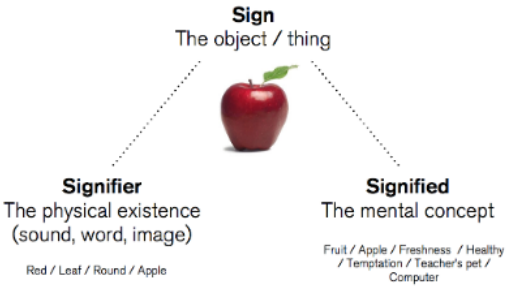


Fig. 5.1 The Semiotic Triangle

The semiotic triangle originally proposed by Saussure in 1907 has been extensively used and developed in linguistics, the arts, and several other fields. Scholars who contributed to its evolution in the domain of music include: Pierce (1893-1913), Langer (1953, 1956), Stefani (1973, 1976), Nattiez (1976, 1987, 1990), Molino (1970-1990), Delalande (1972-1996), Agawu (1991), Windsor (1995, 2000), Roy (1996, 2004), Monelle (1992, 2000, 2006), Turino (1999), Addis (1999), Bayle (1986-2000), Adorno (2002, 2006), Clarke (2003, 2005), Atkinson (2007), Agawu (2008), Kim (2008), Tagg (2013) et al. Although no universal definition of semiotic theory is available within the context of music, Molino, Underwood and Ayrey provide a useful post-structuralist elucidation. They state:

> Molino does not define a set of abstract analytical procedures, but proposes a schematic theory of music as communication, which … facilitates the structural analysis of the "neutral" level without marginalising creative (or "poietic") and interpretive (or "esthesic") factors. These three dimensions, forming the "tripartition", are the cornerstones of a view of music as a symbolic system, in Molino's terms "the symbolic operations or processes that necessitate, if not a communication in the strictest sense of the word, at least a network of exchanges among individuals", common to language, painting, the plastic arts, religion and science. (Molino, Underwood and Ayrey 1990: 105)

To paraphrase, Molino was able to unite structural and post-structural principles by proposing a model that allows signs and symbols (supplied

by the musical materials themselves) to be subjectively interpreted by composers, performers, and listeners, offering the possibility of each arriving at different associative meanings. His tripartite model - discussed in *Chapter One: Modes of Listening* is also relevant within the current context, as in practical terms, it is useful to consider the semiotic triangle *overlaid with* Molino's and Nattiez's (jointly realised) tripartite model - where the *trace* acts as the *signifier*, allowing the musical materials to be interpreted from multiple standpoints. This results in multiple variants of *signs* and that which is *signified*, or in other words, multiple understandings and even divergent meanings delivered through the same musical materials. More recently, a few researchers from the domain of electroacoustic music (Windsor, Atkinson, Kim, et al.) have made valuable contributions to the field in successfully applying both historical and contemporary semiotic theory within the context of listening to artful sound. Luke Windsor re-presents the three types of *signs* originally proposed by Pierce in the early 1900s: icons, indices, and symbols as significant in the process of forming associations with the human body. He states:

> It is clear that indexical and iconic forms of semiosis may be equally important in defining extrinsic reference in music. For example, music may indexically signify the effort of the body that performs it or bear some iconic resemblance to prosody in speech or emotive bodily gestures (see e. g. Clynes; 1977; Lidov, 1987; Clarke, 1993). (Windsor 1995: 60)

Windsor respectively equates Pierce's terms *indexical* and *iconic* with *causal* and *metaphorical* ways of understanding sounds. His thesis is significant in that it offers an alternative strategy for conceptualising the experience of listening to music and sound in assigning two main categories of understanding: sounds perceived to be produced by human agents, and sounds appreciated through their metaphorical associations. These deductions are highly significant to our current discussion, and to the emergence of multimodal, and multisensory understandings of the listening experience. Simon Atkinson (2007) and Daniel Barreiro (2010) also provide excellent summaries of historical and contemporary semiotic sources relevant to the context of the listening experience - with

both researchers pointing out useful semiotic parallels with Smalley's interactive and indicative relationships as well as associations with structural hearing. Barreiro writes:

> Following Atkinson (2007), it is possible to notice that the interactive relationship is related to Jakobson's concept of introversive semiosis, understood as the references established between each sonic event and the other events in a work. Therefore, it is interesting that the interactive listening relationship proposed by Smalley integrates Schaeffer's modes three and four, embracing not only the apprehension of the inner features of the sounds (through reduced listening), but also the musical relationships established between different sonic events (structural hearing). Extroversive semiosis, on the other hand, relates to the indicative relationship and concerns the links with aspects of the world that lay outside the work and the sounds themselves. (Barreiro 2010: 37)

Suk-Jun Kim (2008) extends on Windsor's work in offering *A Quaternary Framework for Electroacoustic Music Listening and Acousmatic Reasoning*. His thesis combines the concept of metaphor (iconic form of semiosis) with an abstract notion of *place.* This provides the four basic categories of *body, place*, and *non-body, non-place* (represented in figure 5.2 below). Interestingly, Kim's retention of the two main categories *semiotic listening* and *spectro-morphological listening* respectively bear a striking resemblance to Smalley's *indicative* and *interactive* relationships.

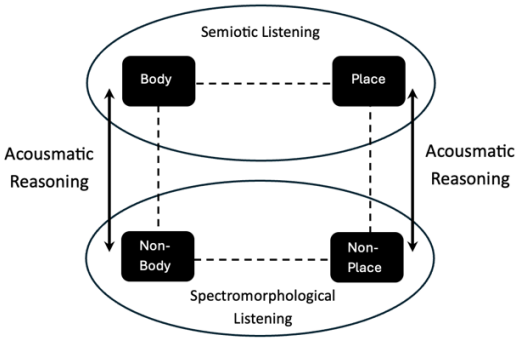


Fig. 5.2 A quaternary listening framework for electroacoustic music based on acousmatic reasoning (Kim 2008: 43)

A quality that unites the work of Windsor, Atkinson and Kim is the way in which their semiotic models cross the boundaries imposed by the siloed nature of listening modes. They present an amalgamated view concerning the primacy of the *human body* in the music and sound-making/listening process and raise the possibility of other universals that may participate in the listening experience – including associations with *metaphor, the environment, structural hearing,* and *space*.

In applying the semiotic triangle in its most general form (pictured in figure 5.1) to our discussion, referential *sound qualities* act as the *signifiers* that present *signs* (objects or events) to listeners, and meaning emerges from the *signified* – the mental concept associated with the sign. In this way, technique becomes the pragmatic means for signifying meaningful ideas. We have already established that all sounds contain these signifiers, so where materials (signifiers) are simply regarded as *abstract* and managed arbitrarily (considered as the artistic medium), associations in the form of *signs* (objects or events) and what they *signify* (concepts, meanings) may be unwittingly presented to an audience. However, where technique is managed with care and rigour, mutually supportive semiotic relationships may be presented that can enhance the quality of the artwork. A simple example of this type of entanglement (between referential and formal modes of listening) might be a stepwise rising melody played on a violin. A compositional process reflecting attention to detail might at least consider the associative (referential) meaning of the rising pattern within the context of the applied formal technique (stepwise rising melody).

No. 32 (1950) - Jackson Pollock

One of my favourite artworks that presents an ideal semiotic relationship between meaning and technique was first revealed to me whilst visiting the Tate Modern Art Museum in London. I entered the main gallery on the second floor and was immediately struck by a wall-sized canvas of Jackson Pollock's *No. 32* (1950) - a work well-known for the innovative technique of dripping paint on a horizontal canvas. The size of the canvas was of critical importance as it at once presented Pollock's expressive gestures to me on an equal scale to that of his bodily gestures. The black paint danced on the cream background and transformed before my eyes into a three-dimensional rendering of abstract movement. I saw with unmistakable clarity how the drip-painting technique had flowed seamlessly out of Pollock's torso and limbs and how this simple yet devastatingly functional technique, presented with such precision, the sequence of movements that had created the artwork in the first place. Each component assumed an important functional role and there was nothing out of place or superfluous to requirements. The insight I received in that moment was both profound and lasting – it revealed to me the sublime artwork that I carried about with me within the mundane physicality of my own human body. The discovery that the movement of my limbs in abstract form was a painting (a dance, a song and a poem) was, in that moment, an epiphany of life-changing proportions.

In considering Pollock's *No. 32* with reference to the semiotic triangle pictured in figure 5.1, it can be deduced that the size and shape of the paint splotches, their presentation as a sequence of motion, together with a basic knowledge of how they were created act as the *signifiers*. They combine to present the artwork (*the sign*), that together with the aid of creative imagination leads to the apprehension of a profound concept *(the signified)* that has already been described as the meaning of the work (in this case of insightful dimensions). It is known from other

sources that Pollock's technique for *No. 32* was developed in response to his popularised comment: 'my concern is with the rhythms of nature… I work inside out, like nature'. I take this to mean that he was focused on pursuing the authenticity of *natural abstract forms* and that for him this included the movements of his own human body. It can also be deduced, with a little imagination, that Pollock's technique - the way in which he applied the drips to the canvas - relates directly to this concern. The primary role of his technique is therefore as a *signifier* that ultimately leads to profound meaning expressed in the artwork (the sign); and without this profound meaning (the signified) the artwork would not have obtained its pre-eminent status. Regarding the artwork (the sign), the paint and canvas are in this case largely utilitarian – they are concrete aspects of the artistic medium itself (the fabric of the artwork). However, there is one notable difference in prescribed norms: the size of the canvas has been expanded to encompass that orbit of human movement equal in size to the far reach of Pollock's limbs around his torso, strengthening the artwork by suggesting an element of intrinsicality within the materials themselves. It also brings to light other possibilities concerning the use of innovative materials (for example, the artist may have chosen to use or imply *body fluid* as a substitute for paint). Alternatively, had he lived in a more modern era, he may have found a way of tracking and imprinting his bodily movement through space on a 3-dimensional canvas (using digital technology). This raises an interesting sub-question: whether the physical qualities of the materials themselves (the signs) should ideally exhibit intrinsic relationships with both techniques (signifiers) and meanings (the signified) or whether it is acceptable develop concrete aspects of the artistic medium arbitrarily.

Image Schemas as Meaning

A collective view pointed out by Windsor, Atkinson and Kim is that *metaphor* plays an active role in the assignation of musical meaning. As Windsor states:

> The important feature of this view is that musical meaning is not merely an attribute of sound but of the metaphorical structures the listener imposes

> upon sound. Hence, within this perspective, if music is in any way mimetic it achieves this through a process of metaphor: the resemblance between music and the world (real or emotional) is not accidental nor inexplicable, nor is it physical. (Windsor 1995: 56-57)

This comment is drawn from a context concerning metaphorical interpretations of abstract music. This is important to note, as referential and semantic materials by comparison (also noted by Windsor) tend to supply *literal* interpretations that are often far more prescriptive. As Smalley has pointed out, sounds become more abstract as the range of possible sources (surrogates) becomes wider, and as Snyder attests to, the widest of all metaphorical interpretations (of musical events) tends to coalesce in five stated categories of image schemas: up and down, centrality, motion-linkage-causation, linearity: pathways and goals, and containment: inside and outside. Without wishing to re-enter a detailed discussion concerning image schemas, within the current context it is proposed that these abstract frameworks primarily operate as metaphorical containers of meaning – offering structural go-betweens - between the more abstract qualities of sound (including less overt referential qualities) and highly generalised forms of associative meaning. An energy and motion trajectory, for example, might be paired with the *pathways and goals* image schema, and the pattern might be further understood (metaphorically) through non-specific associations with a range of other known events that also develop along timelines (for example, *journeys*). Lakoff and Johnson support this assumption in metaphorically stating '*purposes* are destinations [and] *means* are paths' (Lakoff and Johnson 1999: 179). This comparative way of understanding (of one thing in terms of another) reflects one of the most basic ways in which we learn. An important contribution to this topic comes from Gary Kendall, who attests to the operation of a psychological construct he calls the *event schema.* The concept essentially relates to *boundaries* that are automatically created, to house, and make sense of sonic events. He states:

> The event schema aids the listener's process of making sense of what is heard, but it is essentially an abstract container that must be linked with

> other active elements in the listener's mental spaces in the act of forming meaning. Event takes on a sense of meaning when it is bound with specific circumstances (Kendall 2010a: 68)

The theory (originally proposed by Srini Narayanan 1997) is an extremely useful one as it not only provides a container for the five image schemas discussed but offers a bridge to metaphorical ways of understanding sounds. The abstract nature of image schemas is interesting in that it presents the concept of non-specific metaphorical appreciation of sonic events (based on pattern recognition). This type of referential meaning is open to a range of interpretations as previously described. It might also be useful to consider this open concept of referential meaning as one end of a continuum ranging from non-specific to highly specific metaphorical appreciation of sound. At the highly associative end of the continuum, we might find not only pattern recognition, but recognition of specific objects, events, and relationships that play a part in the emergence of *comparative understanding*.

Conceptual Blending as Meaning

In their 2002 book *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities* Fauconnier and Turner present a compelling argument concerning the way in which the human species has come to understand certain principles in terms of others. The paradigm is deeply embedded in language and is evident in many of the words we commonly use to describe everyday life. To define the term and the importance of the concept of *conceptual blending* within the context of artmaking, I refer to the following passage drawn from Gary Kendall's 2010 article *Meaning in Electroacoustic Music and the Everyday Mind.*

> Meaning in electroacoustic music also emerges in the form of conceptual blending (sometimes called conceptual integration) that has most often been studied and discussed in relation to language and reasoning. Fauconnier and Turner (2002) analyse such blends as arising in networks of mental spaces, and they have developed models for the common forms of networks called conceptual integration networks. The prototype integration network has four spaces. Two represent the mental spaces for the input elements that are being blended. One represents the space of the

> common elements that relate the two inputs, the generic space. Finally, selective content of the input spaces is projected into a new space that represents the blend, which now contains relationships that cannot be represented in either input. (Kendall 2010a: 70)

A graphical representation of these four spaces is offered in figure 5.3 followed by examples to illustrate the wide-reaching application of the concept.

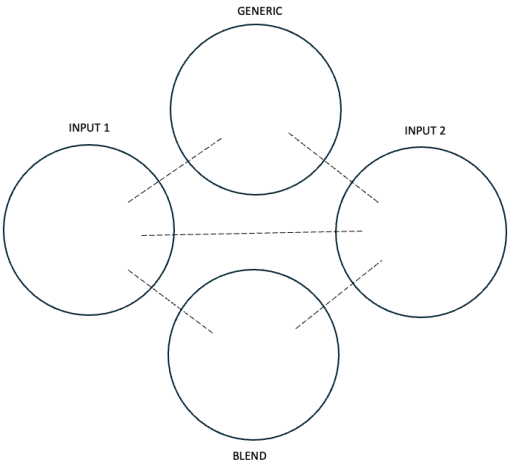


Fig. 5.3: A connection of mental spaces representing a prototype integration network after Fauconnier and Turner 2002

Input 1, for example, might be the sound of a person breathing, While input 2, might be the sound of a wave breaking. A common attribute is immediately identified (in this case the ASR envelope of both sounds). A blend is also produced – a sort of hybridised wave-breath. To this established concept, we may add complexity in the form of internal relationships that may exist within each of the input spaces. Aristotle proposed a related *formula* as a means of constructing such metaphors, and as an illustration of the way in which humans tend to learn. His equation (A is to B as X is to Y) will be used here as a means of further

explaining and populating the *conceptual blending* model. In identifying internal relationships within both input sounds (breathing is the body (A is to B), as waves are to the shore (X is to Y), the metaphor becomes more complex. Aristotle pointed out that the metaphor is produced by switching B and Y (resulting in the equation A is to Y as X is to B) - in this case 'breathing is to the shore as waves are to the body'. The resulting metaphors are 'breathing-shore' and 'wave-body'. As A-B and X-Y relationships become closer, more common aspects become highlighted. Conversely, as A-B and X-Y relationships become more remote, more tentative associations become highlighted. The example 'wind is to fire, as distance is to love' illustrates such a remote association. In this case the explanation 'it fans the great and extinguishes the weak' is normally required to identify the commonality (the generic quality). In applying Aristotle's equation, we produce 'wind-love' and 'distance-fire' – two remote, yet highly descriptive metaphors/conceptual blends.

Image schemas (up and down, centrality, motion- linkage-causation, linearity: pathways and goals, and containment: inside and outside) can assist in the process of identifying these complex internal (A-B and X-Y) relationships. The containment schema is particularly useful, as the previous example '*waves are to the shore*' illustrates. In this case the waves are contained by the shore; however, the wave may also be conceptualised as the container, as in the example *surfer is to wave.* The resulting equation would therefore be *surfer is to wave as breathing is to the body,* producing the metaphors *breath-surfer* and *wave-body*. This mental act of conceptualising and fabricating such relationships is useful as a means of identifying the metaphorical *potentials* of materials through the lenses of all eight modes of listening identified. Such relationships might be conceptualised as *referential blends* that may also be deliberately employed as integrative structural devices (this topic is addressed in detail in *Chapter Ten: Structural Concepts*).

In the context of acousmatic music, *conceptual blending* 'works' primarily through referential qualities that make themselves apparent during the listening process; however, the assignment of at least two component

associations (A-B and X-Y) is required for the method to function at full capacity. This does not present issues when creating music that contains overt references and/or semantic materials, but in the context of abstract music, as has been pointed out (in chapter four), it often falls to *expression* and/or *image schemas* (pattern recognition) to produce meaningful discourse. There is an obvious imbalance between the previous example of *the breathing shore* and an abstract musical work for [say] flute and piano where meaning may be primarily derived from tonal centrality and energy and motion trajectories. Within these contexts, a full analysis of musical materials in terms of the eight modes of listening identified (expressive, formal, referential, visual, gestural, environmental, spatial, and structural) may be useful in identifying such relationships.

Meaning and the Five Layers of Mental Activity

It is important to cross-reference what has been discussed so far as contributors to the assignment of referential meaning with the *five layers of mental activity* model first presented in *Chapter Two, Multimodal Listening* (Fig 2.3). The five respective layers: sensations, gist, locus, contexts, and domains relate loosely to the early and late stages of auditory attention; however, correspondences between all five layers (which are simultaneously active) tend to steer interpretations made at all levels. Perceived associations with gestural agency at the level of *locus*, might, for example, encourage deeper probing at the level of *sensations*. It is also important to iterate that referential meaning is only *one* of the types of meaning identified, and it is likely that multiple modes of listening will be in play within the five mental layers producing a range of meanings. Multisensory relationships and associations may also impact on this process, leading to a diverse range of meanings gleaned from an identical listening experience. Gary Kendall provides a summary of this common experience as follows:

> In the moment-to-moment flow of listening, just as in everyday life, the mind is constantly engaged in determining the meaning of things. A practical question that we are often implicitly posing is: What is going on? That question is most easily answered in terms of a limited number of options. The more that the details can be related to known models and

> categories, the easier it is to have a manageable experience. Meaning emerges from the web of relationships we construct, especially between the onrush of sensory stimulation and the systematization of internal representation. We expect different listeners to experience different meanings. Gist, which is most closely related to sensations, is the most likely layer to produce similar meaning across all listeners. Context and domains which are most closely ties to internal representations, are the most likely to produce different meanings, especially connections between novices and experts (Kendall 2016: 39-40).

Referential Artmaking Strategies

Is it too prescriptive to state that a composer's management of musical meaning should be purposeful? Some composers appear to rely heavily on instinct to create their music. They intuit potential emotional and other meaning-bearing effects on an audience rather than using musical materials for their specific meaning-bearing functions. Other theoretical perspectives might allow a piece to speak for itself as a conglomeration of technique and meaning shaped by the multiple socio-cultural qualities it may have attracted during its haphazard process of coming to fruition. Referential meaning, for some, is thrown to the wind, while for others it is central to their way of working. Despite these differing perspectives, a guiding tenet for this discussion is that most composers and sonic artists feel that they have *something to say* through their music. They tend to make predictions concerning specific responses that they hope to evoke in listeners. For some, these concerns might be focused on music-as-expression/emotion or music-as-abstract-code rather than music-as-referential-meaning; however, as discussed earlier, these fields tend to be entangled with one another. Inherent in these assumptions is the notion that music and sound can deliver at least emotions, abstract musical codes *and* referential meanings to an audience through a sonic language that is understood or intuited by both composers and listeners alike. However, as previously noted, the different types of meaning do not present an equal playing field in terms of the reliability of predicted outcomes. Abstract materials tend to result in a wide variety of meanings and are often entangled with emotions, whereas highly referential and semantic materials are generally far more prescriptive – and in extreme cases can

present universal associations (singular meanings). It is the predictability of this category of *referential* sounds that gives rise to a range of effective meaning-bearing techniques that may be employed (or avoided) by composers and sonic artists with at least some sense of surety.

For composers and sonic artists, choosing an appropriate technique to present their unique ideas can be challenging. Although *new* techniques and technologies may appear to be exciting and innovative, they can lack durability and often present unforeseen types of associations. However, techniques fulfil a practical compositional need. They respond to the question – what do I do with these materials? How do I present and develop them to the best of my ability? Simon Emmerson (1986) answers directly in identifying two main methods of managing technique he terms *abstract* and *abstracted syntax*. This section of the chapter is dedicated to the concept of *abstracted syntax* with a focus on the *referential mode of listening* (where referential meaning, supplied through the source materials, may be seen to justify the technique(s) employed). Three intersecting techniques have been identified: *augmentation of meaning-based qualities, justification of technique,* and *conceptual blending as meaning*. Each category is extremely broad in its scope and can accommodate a wide variety of meaningful (referential) ideas. The techniques may be used separately, or in combination with one another, or may be avoided altogether. A uniting quality is their reliance on referential meaning to form the basis of their communicative functionality.

Augmentation of Meaning-Based Qualities

The key concept behind this identified compositional technique concerns augmentation or diminution of *referential meaning* through one or more of the eight entangled modes of listening identified: expressive, formal, referential, visual, gestural, environmental, spatial, and structural. While the multimodal and multisensory nature of the listening process ensures that sonic qualities may not be pinned down to any one mode, it has been observed that certain applied techniques often result in specific *modal weightings,* with each maintaining a strand of, or link to referential

meaning. Such relationships might be conceptualised as *referential blends.* To clarify, the human ability to focus attention on specific sonic characteristics, allows each of the modes of listening to be attended in isolation and/or as two-way or multiple blends steered by one or more guiding concepts (in this case referential). Numerous examples of the ways in which the eight identified modes of listening are imbued with referential associations are presented below. The approach also seeks to re-categorise Smalley's nine indicative fields (for reasons that will be made apparent).

The Expressive Mode of Listening

As suggested in chapter three, *expression* is often communicated through nuance-level changes in pitch and durational characteristics (formal qualities) or through narrative (referential/semantic qualities). An upward rising pitch towards the end of a phrase, for example, can suggest uncertainty, whereas falling glissandi might suggest sadness and associations with nurturing. In the time domain, an increase in rhythm towards the end of a phrase might suggest flippancy, while a decrease (slowing down) might suggest seriousness. These specific changes in nuance-level *formal* qualities also bear obvious relationships to *gesture* and the human voice. Such patterns tend to be drawn from everyday speech but may also be used deliberately in abstract musical settings for expressive purposes. Smalley (1996) recognised this field (*utterance*) for its natural association with the *reflexive* relationship (referring to emotion-based likes/dislikes). It also delivers strong referential meanings in the ways indicated. The given examples might be best considered as blends of expressive, formal, referential and gestural modes of listening.

The Formal Mode of Listening

Embracing Smalley's (1996) interactive relationship, as well as the indicative fields *energy* and *motion,* the formal qualities of music and sound have traditionally been considered abstract; however, a twenty-first century view would insist that even these sounds present referential meanings. In his 2007 book *Living Electronic Music* Emmerson writes:

> Interestingly, as Kofi Agawu has emphasized, there are plenty of extra-musical references in what is often considered "abstract" classical music (Agawu 1991). For example, we might give "trumpet or horn call motif" = military or hunting (even when played on the piano) as an obvious case, but such a relationship might be extended to the timbral domain of the kind trumpet or horn *sound* = military or hunting. (Emmerson 2007: 6)

Such a description speaks directly to the simultaneous activation of *formal* and *referential* modes of listening that might also be described as *formal-referential blends*. Furthermore, referential qualities have been pointed out as existing even at the highest levels of musical abstraction where patterns may be represented as *image schemas.* Here, *spatial* associations create metaphorical relationships that support referential meanings. For example, Lakoff and Johnson supply highly specify metaphorical assignments to various types of movement (previously presented in chapter four). Their research specifies at least three-way blends between formal, referential, and spatial modes of listening.

The Referential Mode of Listening

The current discussion concerns associative meaning in terms of *thingness.* In highlighting the referential mode of listening in isolation, we are temporarily suspending the principle of modal blending for the purpose of focussing on referential meaning presented through the spectromorphology of sounds alone. The category absorbs Smalley's (1996) indicative relationship, specifically highlighting the indicative field *object/substance.* The domain of acousmatic music is unique in that associations with known physical objects (including the full range of musical instruments) is regarded as a variable (spectral) musical parameter rather than an actuality fixed by the physical constants of universal laws. This gives rise to several powerful techniques including spectral *metamorphosis* between sounds – the purpose of which may be to *suggest* certain associations with meaning-bearing objects, and/or as a means of managing structural passaging between component structures (for structural integration purposes). Timbral characteristics of this type are often extremely well rendered through sound alone supplying the '*what'* qualities of referential meaning.

The Visual Mode of Listening

In *Chapter Six: Audiovisual Relationships*, it will be discussed how this important mode of listening can dominate the listening experience - even in what has traditionally been considered sound-only settings. Absorbing Smalley's (1996) indicative field *vision,* the category presents the concept of *visual blends* with the referential mode of listening. Visual images (whether real or imagined), are by nature, highly associative, and can supply much in terms of referential meaning. For example, it is well-known that live visual images of musical performers actions can dominate the listener's understanding of musical meaning (this topic is also explored in detail in *Chapter Seven: Gestural Associations*).

The Gestural Mode of Listening

The *gestural mode of listening* proposed here absorbs Smalley's (1996) *gesture, utterance*, and *behaviour* fields (although it is limited to the human elements of his behaviour field). The term is used as a broad catch-all for all associations with human agency. Almost any sound may be *humanised* through a process of synthesis or resynthesis that augments its *gestural* content. In this way, referential meaning is formed through a perceived relationship between a performer (real or imagined) and the sounding object or event. From a practical standpoint, *augmenting* these qualities through digital production processes might include convolving a gestural sound with data from its own amplitude and frequency analysis. It may also incorporate performing (or re-performing) pitch, time, volume, and perhaps spatial fluctuations using gestural controllers. This type of sonic transformation also evokes at least expressive, formal, and spatial modes of listening. Transformations in the time domain also provide a bridge between gestural (performative) and textural (environmental) domains. These naturally occurring *multiple blends* (between numerous modes of listening) tend to be heavily imbued with referential meanings.

The Environmental Mode of Listening

Common use of the term *landscape* within the context of acousmatic music is evidence enough of the entanglement between environmental,

visual, spatial, and referential modes of listening. Trevor Wishart provides an early but useful definition, which, at that time was also an attempt to *redefine* the term *landscape* within sonic contexts. He states:

> We may effectively break down our perception of landscape into three components which are not, however, entirely independent of one another. These are: I, the nature of the perceived acoustic space; II, the disposition of sound-objects within the space; and III, the recognition of individual sound objects. (Wishart 1986: 45)

Wishart then goes on to propose metamorphosis of landscape/soundscape components to produce entirely convincing, yet *otherworldly* environmental spaces. The associative qualities of these environments tend to mean different things to different people, although universals may also be observed (such as environments that are conducive to human life). The category absorbs Smalley's (1996) indicative field *environment,* as well as Truax's (1984) three modes of listening (background listening, listening-in-readiness, and listening-in-search) and in this context specifically highlights environmental-referential relationships that will be discussed further in *Chapter Eight: Environmental Connections.*

The Spatial Mode of Listening

Encompassing Smalley's (1996) indicative field *space* as well as his elaboration of the field offered in 2007, the category presents several pragmatic techniques relating to spatial composition - especially within the context of multichannel acousmatic music. These include ambisonic recording and decoding techniques that evoke bodily sensations of *envelopment* (being there), associations with energy and motion through spatial *pathways* presented through multichannel loudspeaker arrays, and methods of transformation and/or metamorphosis *between spaces*. Even these abstract spaces (small/big, up/down, etc.) cannot be divorced from their referential associations. They can *mean* a great deal and form an important part of the listening experience, and in this context might be considered spatial-referential blends. These spatial relationships will be examined further in *Chapter Nine: Spatial Forms.*

The Structural Mode of Listening

Often categorised as a sub-field within formal musical qualities, the structural mode of listening is primarily concerned with the organisation of sounds in terms of their component structures, and/or structural integration. Specific techniques may also form natural associations with *image schemas*. One such example is the emergence of a textural trajectory where one speaking voice (expressive, referential, gestural) increases to many over time. The construct evokes the pathways and goals image schema, where – somewhere along the trajectory path, well before the climax is reached, the listener can predict a natural endpoint (in this case perhaps thousands of simultaneous voices). The example reveals the creative potential to either fulfil or thwart those expectations, and these patterns often carry with them referential associations. For example, an expectation *thwarted* can present quasi-narratives relating to prevention/blocking or unexpected discovery. Such structural-referential relationships are also at the heart of the technique of conceptual blending, which will be discussed in *Chapter Ten: Structural Concepts.*

In summary of this section, the eight entangled modes of listening identified emphasises that the assignment of meaning (in general) may not be pinned down to specific sonic qualities derived from the *referential* mode of listening alone. Rather, it arrives through a network of *affordances* provided through the eight identified modes of listening identified operating in parallel. From the several examples provided, it may be observed that one mode of listening cannot survive in the absence of others. They all tend to work together to provide *expressive, formal, referential, visual, gestural, environmental, spatial, and structural meaning.* Seen through the multimodal, multisensory lens (or lenses), these fields may be seen to operate simultaneously through the combination of multiple listening modes and/or also as isolated contributing fields in their own rite.

Justification of Technique

The example of Jackson Pollock's No. 32 outlined above illustrates that the techniques we choose have implications, foreseen and unforeseen. They *mean* certain things and therefore should not be used arbitrary. Justifications for the use of identified techniques (for example, drip painting) will invariably be *sought out* by listeners. It is also noted that any technique that is overt or overused, tends to require additional justification. In the absence of supplied justifications (supplied by the artmaker), the receivers of art and music tend to formulate their own perceived reasons for the use of such techniques. This important relationship has already been highlighted through the semiotic triangle presented in figure 5.1. Within that context, technique may be simply defined as *the way in which* signifiers are presented. To expose the creative potential of this relationship, the following scenario is examined. A sonic artist working in a multichannel acousmatic studio decides to pan abstract materials (an undefined noise) around their eight-channel circular loudspeaker array. The effect is sensationally pleasing, however, from the perspective of referential meaning, it has not yet acquired its semiotic concept. The composer then reflects on *objects that go around*, (mechanical actions, cycles of the sun and moon, maypoles, etc.), and decides on an appropriate concept that not only becomes available for discovery during the ordinary listening process (justifying the technique) but also gives rise to a more sophisticated range of associated *surround panning* techniques. The observation presents the possibility of engineering such precise and intact relationships reflecting the concept of *justification of technique*. Four further examples have been chosen to illustrate this potential, which can be a powerful method of proceeding for composers and sonic artists.

The first example is weighted towards the mode of listening concerning *formal musical qualities.* This culturally supplied frame, making use of abstract characteristics such as pitch, rhythm, and duration, is often dictated by the constraints of the musical language itself, providing a semblance of commonly observed developmental possibilities supplied by the listeners' understanding of the genre employed and the range of

musical choices available within it. This type of justification invariably involves a process of pattern recognition – identified first within the music and then compared with other previously recognised patterns (previously heard in the work or other works and stored in long-term memory). A rising melodic pattern for example (up/down image schema), might justifiably continue to rise in reoccurring episodes, develop to include variations, or be transferred to other instrumental textures. Such developments *make sense* as they are justified by the materials themselves. Management of formal musical qualities in this way adheres to the ideal that some semblance of justification (in the form of shared patterns) might be maintained even the most abstract of musical contexts.

The second example makes use of a combination of *formal* and *referential* modes of listening to justify techniques. The method is one that is used prolifically throughout instrumental/vocal composition and sonic arts. The *violin-ness* of a sound, for example, supplies the justification for an extensive range of bowing, plucking, scraping and hitting techniques, although interestingly, the more extended the techniques are (how far they depart from known and accepted playing techniques) the more they tend to demand additional justification. A hallmark of this category is the containment of justification within the spectral signature of the materials themselves, where the multimodal nature of sound can present both referential, formal, and other qualities to the listener simultaneously. In the given example, the violin brings with it a sociocultural framework and history within which certain activities (techniques) are acceptable to an audience whilst others are not. The recognisable spectrum of the violin sound itself is enough to supply this *frame*, providing justification for certain techniques (and not others) to be carried out.

The third example involves *semantic justification* – a subset of the *referential* mode of listening. The technique is characterised through the concept of *narrative*. Recognisable speech is an obvious example from this category, but other materials too, such as music and sound from historical events can provide a network of signs powerful enough to bring to the fore ideas or stories that heavily steer the listening process. An overt example

from this category (previously used) is the continuous repetition of the prerecorded vocal phrase 'loop this phrase'. Considering this example alongside a sound that is arbitrarily looped (that is, one that contains no such justification) it is easy to ascertain the precise means by which the listener self-assigns just cause in both instances. In the first case, the semantic quality of the vocal phrase supplies a code (through the English language) that justifies the ensuing technique (the phrase is indefinitely looped). In the second case, in the absence of such intrinsic justification, the listener is compelled to self-assign the most satisfactory form of just cause available. Thus, for most listeners, the loop becomes an artifact of the technological process itself (exposing the technique as a contrivance) or may develop loose associations with other things '*looped*' such as drum and bass music or mechanical actions.

The final example features a combination of *environmental* and *spatial* modes of listening, adopting principles of Newtonian law for the purpose of justifying techniques. For example, the *energy and motion* qualities of the sound of a bouncing ball might be extended to include friction, gravity, resilience (bounciness), three-dimensional spatial position, as well as collisions with other sounds. Observing these natural laws of physics (for example, the principle of continuity), tends to supply listeners with highly plausible (justified) pathways for development of materials. It is the deliberate and/or intuitive use of these known laws and schemas, observed in nature and accepted as constants through life experience, that often results in such convincing musical outcomes.

The human need to *make sense* of actions (and in this case musical development) presents a general rule: that *transformation of musical materials should ideally not be carried out arbitrarily as any development will invariably make a comment on the original materials themselves*. Returning to the question posed at the end of the Pollock analysis:

- Whether the physical qualities of the materials themselves (the signs) should ideally exhibit intrinsic relationships with both techniques (signifiers) and

meanings (the signified) or whether it is acceptable develop concrete aspects of the artistic medium arbitrarily.

Our discussion has now established the clear view that materials should not (ideally) be used or developed arbitrarily - justification for the use of specific (particularly overt) techniques might ideally be supplied.

The Breaking of the Scream (1999) - José Halac

I would now like to illustrate the uniting potential of the concept of *justification of technique* using José Halac's *the Breaking of the Scream* (1999) as an exemplar. The stereo acousmatic work begins with a scream of nearly 60 seconds in duration then dovetails into a 10-minute work presenting a constant vocal texture. During the listening process, information is presented encouraging the listener to seek out a justification for the form and structure of the work, which is unrelenting in its pace and pressure of delivery. The insight - that the entire piece is in fact a *single scream* arrives as an epiphany, uniting the materials structurally and retrospectively justifying the techniques employed. The profound discovery that *all* sounds within the work represent different aspects of human emotion contained within a single scream effectively transforms what was previously regarded as a relatively banal human utterance into a multifaceted kaleidoscope of musical expression approaching the sublime. On making this realisation, the title of the work *The Breaking of the Scream* also cleverly reveals itself as a play on words. A hallmark of this ideal species of *meaningful creative idea* is observed whenever the discoverable meaning contains conceptual information that may thereafter be used to justify the techniques employed.

As part of the process of preparing this chapter, I asked the composer (José Halac) to comment on his creative approach to the work. His response directly confirms the deliberate employment of this stated ideal:

> I thought of putting EVERY sound in one passage, obscured by reverbs and granulation ambience out of which the scream must emerge. So, the scream is actually a convolution of all that - and will gradually compose the piece. So out of it, comes the piece. Like a seed which contains everything. (personal email 19:5:21)

It is important to note that this is *one* of the meanings that might be gleaned from the work. Through other communications with Professor Halac it is understood that, for him, the work speaks on multiple levels. The range of techniques employed in the piece deliberately embrace the expressive, formal, and referential/semantic modes of listening but also correspond to the remaining five multisensory modes: visual, gestural, environmental, spatial, and structural. It would be fair to say that his listening experience is the most detailed and profound of all listeners, but this does not preclude other listeners from making their own unique discoveries and deductions that the composer may not even be aware of. Interestingly, this capacity for divergent interpretation does not diminish the quality of the work. To the contrary - it brings to the fore the striking variety of profound insights that become available to individual listeners through the realisation of a work of art that attends to the presentation of both referential meaning and technique with such care and rigour.

Conceptual Blending as Meaning

The reliance on *metaphor* in our daily lives, as a method of understanding one thing in terms of another, has already been discussed. The paradigm extends throughout language, is embedded in our archetypal spatial constructs, and forms an important strand in terms of the assignation of referential meaning in musical contexts. The art domain in general is one that allows us to compare one thing in terms of another in ways that challenge our understandings and provide new insights from previously unimagined perspectives. This technique termed *conceptual blending* – has been knowingly and unwittingly adopted by thousands of artists the world over. Comparative statements based on Aristotle's formula (as a central concept) bring into our consciousness aspects of *referential meaning* that are, within these contexts, at the core of artistic communicability.

While the Sun Shines (2016) – John Coulter

To further illuminate the technique of *the conceptual blending* within the context of acousmatic music, I refer to an excerpt from *While the Sun Shines* (Coulter 2016). In the extract, a senior male voice (my father) describes the mechanical sounds of a hay baler through a series of skillful vocal imitations followed by the spoken phrases 'and those were the knots being tied … and you'd listen for every knot'. The referential qualities of the materials themselves – that is, the *baler-like qualities* of the vocal imitations - provide justification for the ensuing technique, which is to metamorphose his speech into the actual sounds of the mechanical baler. Although the transformation is overt, it is accepted by the listener due to this in-sound method of justification; however, the metamorphosis of man-to-baler is, at this point, lacking in meaning, which forms an important question in the mind of the listener: *what does this transformation mean?* The answer is not immediately supplied but is suspended to allow a different kind of story to be told – one of family loss and tragedy. It is not until the final moments of the work that the missing information (the meaning) is delivered in the form of a metaphor - the double meaning of 'knots being tied' within the context of both hay-baling and the metaphorical knots of human interconnection, representing the resilience of a farming community in times of grief and loss. The two concepts explicitly stated in the form of Aristotle's formula (A is to B as X is to Y) are *knots are to hay bales* as *people are to a small farming community* – produce the metaphors *people-bales* and *community-knots.* As seemingly remote as these two concepts might first appear, the discovery of the metaphor at the end of the work during the ordinary listening process forces a retrospective recontextualisation of the 18-minute work in its entirety. The technique presents an element of effective artistry on the part of the composer, but more importantly, it structurally unites and cadences the artwork.

In essence, the technique of *Conceptual blending* presents *relationships* that allow the emergence of a *third shared space* that might also be

described as *auditory Gestalten* - a product that is *other than the sum of the parts*. It relies on the *superimposition* of at least two meaning-bearing sounds that may be presented in a variety of ways (for example, simultaneously, or sequentially). Variations of the technique have been identified falling into two main categories: simple and complex.

Simple conceptual blends rely on the superimposition of two basic concepts with reference to the four mental spaces diagram in figure 5.3. Materials may feature several active modes of listening that work in combination to supply various forms of meaning. For example, a young woman speaking about her life (referential/semantic mode) presented simultaneously with Chopin's *Nocturne in C-sharp minor* (expressive, formal modes) may suggest that she is unhappy with her life – and even highlight micro inflections in her voice that support this assumption. The technique may be reduced to a simple overlay of two schemas: *young woman* and *sad music* combining to present *sad young woman.* The example bears a striking resemblance to *music for film* that often employs abstract music for its mood-enhancing potential. Yet another way to conceptualise this example might be to imagine the combination as a metaphor – young woman *as* sad music.

Complex conceptual blends feature internal relationships within both inputs 1 and 2 as represented in figure 5.3. Several examples displaying such relationships have already been supplied with the aid of Aristotle's formula *A is to B as X is to Y;* however, as previously noted, complex conceptual blends are often limited to delivery through highly referential and/or semantic sounds and images. Their relationship with *language* is apparent. It is not so easy to identify these types of internal relationships in abstract materials, although as has been demonstrated, referential qualities are certainly alive and well within all eight of the entangled modes identified. Complex conceptual blends will be discussed in further detail in *Chapter Ten: Structural Concepts* along with principles of auditory Gestalten that are known to impact on the process of structural integration.

Summary and Conclusions

In this chapter we have examined the many contributors to associative meaning in musical contexts with a focus on the assignment of *referential meaning*. An examination of Smalley's (1996) nine indicative fields first reveals the importance of clearly identifying the contributing attributes and potency of each of the fields for the purpose of accurately presenting '*what*' characteristics, and the referential meanings they *afford* to individual listeners. The semantic qualities of sound were then discussed as key contributors to meaning. Here, the use of literary devices such as imagery, symbolism, flashbacks, foreshadowing, allegory, juxtaposition, etc. are made available to the composer in the telling of complex and emotive stories drawn from human life experiences. Semiotics theory is then employed to break down the assignment of meaning into its constituent parts - drawing on the availability of *signifiers* (sonic attributes within the sound) to deliver a series of associative concepts (*the signified*) *through* the presentation of recognisable sound objects and/or events (*signs*). The semiotic studies of Windsor, Atkinson and Kim, also present important associations with the *human body* as well as other universal concepts such as *metaphor, structural hearing,* and *space*. Many of these principles will be examined further in subsequent chapters within the book. To conclude the first part of the chapter, the active role of *metaphor* in the assignment of referential meaning is put forward in combination with Synder's five image schemas: up and down, centrality, motion- linkage-causation, linearity: pathways and goals, and containment: inside and outside. The paradigm then leads to the technique of *conceptual blending, which* presents four abstract spaces: two inputs, a generic space and a blended space. Complexity is then added to the model in the form of internal relationships (using Aristotle's formula *A is to B as X is to Y*), exposing the potential for the presentation of common and remote metaphors.

Part two of the chapter addresses referential meaning-based artmaking strategies through the presentation and discussion of three inter-related practical techniques. Methods concerning *augmentation of meaning-based qualities* are first put forward using the eight entangled modes of

listening identified as the basis of explanation. It is suggested that the technique might be employed for the purpose of supplying or limiting referential meaning in creative musical contexts. The second technique discussed concerns justification of technique, where techniques act as signifiers for meanings, and meanings justify the techniques employed. Examples are then supplied specifying how '*meaning* and *justification of technique'* might ideally be simultaneously supplied. The final technique examined is *conceptual blending as meaning*. Two variants are observed: simple and complex, with the second exhibiting internal relationships in both component input spaces. This final technique is one that has been observed as being in prolific use throughout the contemporary art world. The way in which it operates allows the receivers of art to compare one thing in terms of another in ways that challenge their understandings and provide new insights from previously unimagined perspectives.

The outcome of the chapter has narrow and wide implications. From a narrow perspective, it provides composers and sonic artists with several highly effective concepts and techniques that may be of use in forming specific types of referential meanings. The three inter-related methods specified highlights the possibility of managing the referential qualities of sound, the reciprocal nature of the assignment of meaning and justification of technique, as well as metaphorical understanding (conceptual blending) in musical settings. Of course, there is no requirement for composers/sonic artists to do so, as previously noted, a great deal of eminent music and sonic art functions without any consideration given to referential meaning whatsoever (they rely entirely on other modes of listening for functionality). Such is the complexity of the multimodal listening experience. The chapter also calls to attention the wider definition of *meaning,* which includes at least expressive and referential meaning, as well as other forms of association and functionality brought to bear through the remaining identified modes of listening: formal, visual, gestural, environmental, spatial, and structural.

Chapter Six: Audiovisual Relationships

A modern understanding of the human listening experience invariably includes some form of registration of vision – even if that appreciation is reduced to what is sometimes referred to as a *quasi-visual* experience. Katherine Norman speaks to this multisensory way of appreciating sound and music in the following comment. She writes:

> We insist on visual perception as our window on temporal presence despite the fact that many other perceptions continually invade our consciousness. As a sound impinges on our listening attention, we create an internal visual presence to bring the experience into our perceptual 'now'. For instance, a telephone conversation makes especial demands on our consciousness because, as McLuhan suggests, the auditory image needs to be strengthened by other senses. (Norman 1996: 6)

The term used in our discussion to date, *the visual mode of listening* is intended to embrace these imaginary visual sources (those Norman refers to as *internal visual presence*) as well as real images attended by the visual senses, defining the widest boundaries of the topic. The term *audiovisual* (used in the title of the chapter), emphases an active relationship between these two types of sensory experiences (audio and visual) - where each sensory type might be responsible for delivering specific types of information that combine to form an *audiovisual* multimodal construct. This presents the narrow context of the topic, and the main focus of the chapter. While the terms are similar, their distinction is important, as the first has a relationship with all genres of music and sonic art - adding a distinctly *visual* component to auditory attention (that may include imaginary and real images), while the second is generally limited to multimedia contexts where audio and visual components are presented together (primarily employing real visual images). The chapter first offers a detailed investigation of the *audiovisual mode* of attention before expanding to the more general topic of the *visual mode of listening.*

The 'audiovisual mode' has been met with much interest from scholars such as, Chion (1994), Cook (1998), Kröpfl (2005), et al, who have applied it within the narrow context of music and sound with moving images. Here, audiovisual artists (film makers, and/or those working with digital video/audio) might ideally consider the content of both audio and visual streams and look to the possibility of selecting materials that complement or offset one another. Composers and sonic artists seeking to supplement their works with visual materials should also understand that various combinations of audio and visual materials will inevitably detract from one another and risk unhinging the musical language that they have worked so hard to create. All creative artists working with sounds and images are faced with these potentials and challenges whether they choose to address them or not. Durk Talsma, and Max Mathews succinctly state the opposing principles.

> Many behavioural studies have provided evidence for the hypothesis that integrating visual and auditory stimuli serves the purpose of enhancing perceptual clarity (Stein, Calvert et al.). These results suggest that communication between the visual and auditory brain areas is a highly effective and relatively automatic process (Foxe et al.)' (Talsma, Doty and Waldorff 2006: 679).

and

> I personally find most combined music-video art problematic. It seems to me that the sound and images often compete for my attention… If I pay attention to what I am seeing, I often miss what I am hearing, and if I try to concentrate on the music, the images can often be an irritating distraction… (Mathews 2005: 94).

This chapter seeks to transcend this paradox through the identification of audiovisual relationships that function in a variety of homogeneous and heterogeneous ways. This narrow context (focused on audiovisual media) provides a level of understanding that is also highly relevant in wider contexts (all combinations of audio and visual materials). The author's model for classifying audiovisual relationships (most relevant to digital audiovisual production) published in *Organised Sound* (Coulter 2010) is

first revised and restated, and examples are offered to illustrate aspects of audiovisual *language* put forward. Adam Basanta's 'Audiovisual Relations' model published in *eContact* in 2017 is then cross-referenced, followed by an investigation of Diago Garro's 2005 and 2012 articles in which he states the fundamentals of audiovisual pairing techniques. Finally, a new and improved model employing aspects from all scholars is put forward. This narrow focus of the chapter offers a vocabulary for discussing the functionality of audiovisual materials, detailed methods of identifying audiovisual relationships, and numerous effective techniques of audiovisual production, while the wider findings extend to all associated art domains including the numerous musical genres and approaches to artmaking with sound discussed in this book.

Background

Michel Chion describes the language of sound with moving images as being primarily founded on visual language, referring to 'cinema' as 'a place of images, plus sounds' with sound being 'that which seeks its place'. Here, and in similar settings, he proposes that listeners enter an 'audiovisual contract' where it is the function of sound to 'add value' to the visual images presented (Chion 1994: 68). These difference between the acousmatic (audio-only) and audiovisual modes are also evidenced by empirical studies from the domain of psychology - as Talsma, Doty and Waldorff state:

> It has been found, however, that auditory stimuli are known to capture attention easily (Schroger et al.) and also that the processing of audio stimulus features occurs generally faster than that of visual stimuli (Waldorff et al.) Thus, on the basis of these differences in processing time, we would also predict that attending to the visual modality would affect the multisensory integration process differently than attending to the auditory modality' (Talsma, Doty and Waldorff. 2006: 681).

There are further points to be raised concerning the relationship between 'the acousmatic' and 'the audiovisual'. Firstly, the two modes are quite closely affiliated. For most, the acousmatic includes a *quasi-visual* component (visual relationship), while physically seen and heard

materials constitute the experience of the audiovisual mode. Simon Emmerson elaborates:

> The acousmatic condition excludes clear information on source and cause which we (products of evolution) attempt to 'fill in'. For me (and I believe many others) that process has a visual component. The imagination constructs a quasi-visual mindscape with many of the characteristics of 'real' vision. There are also interesting cross-references to other sense modes – most notably a sense that I might be able to 'feel' textures (Emmerson 2005: 86).

Notwithstanding these similarities, it is also important to note the differences between audiovisual and audio-only modes, for the purpose of effectively managing transitions between them. Chion attests to the absence of a continuum, noting that a collapse of 'the audiovisual structure' is necessary before a transition to the general audio-only mode can occur (Chion 1994: 40). In 2007 a series of tests was carried out at the University of Auckland that confirmed Chion's views. Eighteen postgraduate students took part in the study which investigated the effects of *black screen*, memory residues, minimal images, and colours as transitory states between audiovisual and acousmatic modes. It was found that black screen images tend to evoke a condition like the acousmatic mode, although *mode switching* tended to take some time (ten seconds or more) following complex moving images. Surprisingly, it was also found that minimal images (including colours, partial blacks, etc.) tend to demand audiovisual attention in the same manner as more complex moving images. A notable exception (a *mid-point*) was observed when white screen was used in place of black screen. Most participants experienced sensations of *retaining* the acousmatic mode; however, they also reported that the associated sounds took on *white* characteristics.

A Model for Classifying Audiovisual Relationships (Coulter 2010)

An important concept put forward by Denis Smalley in 1986 concerns 'the dual potential of sound' (this refers directly to the abstract and referential qualities of sound represented by the *formal* and *referential* modes of listening). While this principle has now largely been supplanted by

multimodal perspectives on the listening experience, in 2010, the concept was used as the foundation for both audio and visual categorisation in the author's 'model for classifying audiovisual relationships' published in *Organised Sound* (presented in figure 6.1).

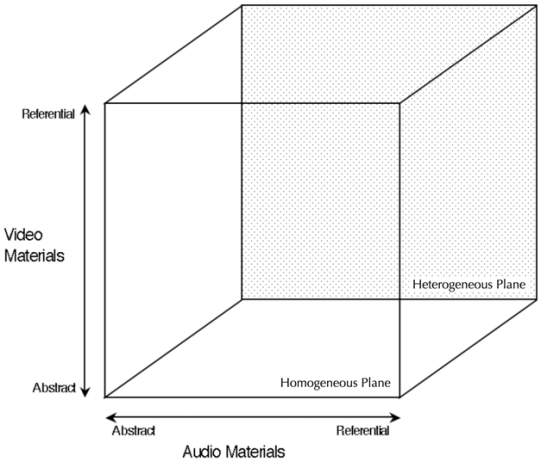


Fig. 6.1: A Model for Classifying Audiovisual Relationships (Coulter 2010)

Within the context of attending *audiovisual* materials, the intersection of referential/abstract sounds and images yields four groupings: referential video-referential audio, referential video-abstract audio, abstract video-referential audio, and abstract video-abstract audio. These relationships form automatically, as it is known through psychological studies that humans have a natural tendency to integrate audio and visual component textures - as Francisco Kröpfl states:

> Even in extreme cases of heterogeneity between both dimensions [the] mind needs to establish a coherent link in a process of simultaneity, thus, in spite of a non-existent will of organisation a structural integration is inevitable. (Kröpfl 2005: 90).

However, experience tells us that we are also capable of separating the two streams of information. This allows each media type (audio and

visual) to be evaluated independently, and, in remote circumstances, for audio and visual components to be attended as dissociated information stems. An example of this phenomenon is driving a car while talking on a cell phone. Interestingly, in situations that require such divided attention, cognitive ability also tends to become *bottlenecked* – limiting capacity to fully attend both streams simultaneously. The *Necker cube* in figure 6.1 provides an established construct for considering the natural tendency to integrate, as well as the voluntary ability to separate audiovisual materials. This is achieved visually by concentrating on the shaded heterogeneous plane until it is perceived as a foreground image.

Cook's Three Models of Multimedia

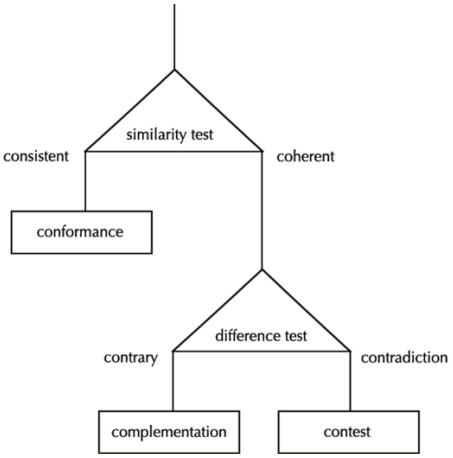


Fig. 6.2: Cook's models of conformance, complementation, and contest.

In his 1998 book *Analysing Musical Multimedia* Nicholas Cook proposed a similar model for identifying three different types of audio and visual media pairs that constitute the audiovisual listening-seeing experience. He writes:

> Figure [6.2] represents the relationships between what I see as the three basic models of multimedia, which I am calling conformance,

> complementation, and contest. I have characterised multimedia as predicated on a distinctive combination of similarity and difference; so it is logical that the three models should relate through what I call the similarity test and the difference test (shown by the triangles) … Conformance begins with ordinary meaning, whether located within one medium or diffused between all; contest, on the other hand, ends in meaning. And as the association of conformant models with synaesthetic and metaphysical speculation demonstrates, conformance tends towards the static and the essentialized, whereas contest is intrinsically dynamic and contextual. The mid-point between these two extremes is represented by the third model of multimedia, complementation, which fig. [6.2] represents in negative terms as that which exhibits neither consistency nor contradiction (Cook 1998: 98,103).

There are two main points of difference between Cook's model (figure 6.2) and the classification of audiovisual relationships model (figure 6.1) that are worthy of mention. The first relates to the similarity and difference tests proposed by Cook. The distinction is that, in the 2010 model, a more sophisticated test is applied - resulting in eight possible outcomes (Cook's model provides three). The second point relates to Cook's inclusion of his third (mid-point) model. In figure 6.1, the category is included as part of a continuum that operates between the extremities. As such, it may be helpful to consider the two models as complementary (with the 2010 model extending on Cook's earlier views).

Cohen's Eight Functions of Music in Multimedia Contexts

Annabel Cohen adds complexity to the discussion through her proposal of *eight functions of music within film or multimedia contexts*. A few of the functions described by Cohen are adequately represented by Cook's model (which is based on modes of listening/seeing), while others (such as leitmotif) are more effectively explained in terms of their historical genesis or status as established production techniques.

> Cohen (1999a) described eight functions of music in a film or multimedia context. First, Music Masks extraneous noises. Second, it provides continuity between shots, for example, when the camera alternates between close-ups of two people who are presumably looking as each other… Third, … it directs attention to important features of the screen

> through structural or associationist congruence. Fourth, when unassociated with a particular focus it induces mood, as often occurs during the opening credits of a film… Fifth, it communicates meaning and furthers the narrative, especially in ambiguous situations… Sixth, through association in memory, music becomes integrated with the film… and enables the symbolisation of past and future events through the technique of *leitmotiv*… Seventh, music heightens the sense of reality of or absorption in film, perhaps by augmenting arousal, and increasing attention to the entire film context and inattention to everything else… Finally, music as an art form adds to the aesthetic effect of the film. (Cohen 2001: 253)

Within the current context of classifying audiovisual relationships, functions three and four (associationist congruence and mood) present equivalent concepts to the homogeneous and heterogeneous principles proposed in figure 6.1. The six remaining functions might therefore be considered as specific techniques that are available to audiovisual artists relevant to the production of mixed-media materials. It is also worth noting that Cohen's approach appears to be closely aligned with traditional film scoring techniques and may therefore begin to lose relevance in wider audiovisual artmaking contexts.

Concomitant and Isomorphic Relationships

Extensive tests by psychologists such as Tasma, Doty, Waldorff, et al have confirmed that integration occurs involuntarily when congruous relationships are perceived to exist between audio and video materials. This general law (supplied by nature) is at the heart of two types of audiovisual relationships that have been observed: 'concomitant' - defined as 'going together, accompanying, concurrent, attendant'; and 'isomorphic' - defined as 'corresponding or similar in form and relations.' (Oxford English Dictionary 3rd Ed.). In settings where audio and visual information is divided (for example, sitting around a fire, and hearing a noise from the surrounding bush) audio and visual streams tend to carry different information. In other settings audio and visual streams can be seen to carry the same information (for example the striking of a hammer). The latter type of audiovisual relationship is reinforcing, and psychologists (such as Tasma, Doty, Waldorff, et al) have shown that super-additive effects occur through integration of such commensurate

audio and visual materials (these effects will be discussed later in the chapter). Division of audiovisual relationships into these two main categories is also supported by the research of Chion (1994), Cook (1998), Cohen (1999a), and Kröpfl (2005).

Audiovisual materials exhibiting *concomitant* relationships (heterogeneity) are integrated through a natural process of highlighting and masking. The process results in a hierarchy of textures - the greatest of which have the potential to form functional materials within the work, and the least of which are entirely ignored. Chion refers to this phenomenon as 'synchresis'. He states: 'Play a stream of random audio and visual events, and you will find that certain ones will come together through synchresis, and other combinations will not' (Chion 1994: 65). Audiovisual materials displaying *isomorphic* relationships (supporting homogeneity) contain features that act as catalysts in the process of audiovisual integration. These qualities include shared formal characteristics such as duration, intensity, proximity, etc (abstractions), as well as shared meanings (primarily referential). These two qualities are also presented in the dictionary definition of the term isomorphic, 'corresponding or similar in *form* and *relations*.' When both form and meaning are aligned, we experience highlighting without masking. Nature is the architect of intact isomorphic relationships. Often the visualised and sounding result of kinetic energy, they occur spontaneously in our everyday lives and are governed by laws of the natural universe.

In essence, materials that exhibit concomitant relationships rely on highlighting and masking that occurs when two (or more) schemas are simultaneously activated (super-imposed), while materials exhibiting isomorphic relationships rely on the activation of a single common schema. However, it is important to note that this binary description is an oversimplification. In real-world scenarios, audiovisual materials exhibit instantaneous combinations of concomitant *and* isomorphic relationships, and exactly how the combinations affect the process of integration varies greatly amongst listeners/viewers. In considering the two types of relationships (concomitant and isomorphic) within the

context of the four classes of audiovisual relationships already identified (referential and abstract audio and video materials), eight different possibilities are arrived at (one for each corner of the Necker cube presented in figure 6.1). Each of the eight sub-relationships exhibits a unique quality of audiovisual language that may be exploited by composers, sonic artists, and audiovisual artists alike.

Concomitant Relationships

The first species of concomitant relationship concerns referential video and referential audio. Common referential features are highlighted while other qualities less relevant to the process of integration become masked. An example of this type of relationship might be the simultaneous presentation of moving images of [say] a scenic countryside with the soundtrack of a narrator engaged in an unrelated topic. Interestingly, within these contexts the phenomenon of 'synchresis' (Chion 1994) tends to insist on transitory moments of integration even in the total absence of composed 'structural or associationist congruence'.

The second type of concomitant relationship concerns referential video and abstract audio. This category relies on the principle of video-dominant 'primacy of reference' as referred to by both Nicholas Cook and Durk Talsma. Firstly, Cook states: 'Unitary conformance, by contrast, means that one medium predominates, and that other media conform to this.' (Cook 1998: 101). Talsma, Doty and Waldorff echo Cook's views: 'Audiovisual integration processes appear to associate the visual and auditory stimulus components with each other, even when only the visual component was relevant' (Talsma, Doty and Waldorff. 2006: 686). Such is the dominant power of referential video - that accompanying abstract sounds are often transformed in the minds of receivers to align with schemas that are more suitable for integration with video materials. An example of this type of relationship might be a visual montage of the images of war, together with an abstract musical soundtrack. Messages presented by the music will tend to be transformed to align and integrate with the visual materials.

A third kind of concomitant relationship concerns abstract video and referential audio. Again, the category relies on the principle of 'primacy of reference' (in this case audio-dominant). For example, where speech is used as the audio stream and a series of changing colours is presented visually, various components of abstract video textures become highlighted (and given meaning) by specific words, while other qualities less relevant to the process of integration tend to be masked. Chion supports this two-way phenomenon. He states: 'synchresis' occurs 'when vision affects sound and *vice versa*' (Chion 1994:28). In this example, although the audio narrative takes the lion's share of overall meaning, specific words/moments also tend to become adapted (highlighted) in part by the abstract visual images that accompany them.

The final category of concomitant relationship concerns abstract video and abstract audio. This category relies on the intersection of abstract 'images' created by both audio and video textures. This relationship is the most amorphous of the eight identified as it relies on transitory moments of structural and/or associationist congruence that may be experienced quite differently (subjectively) by individual listeners/viewers. An example of materials operating in this way might be the stretched (abstracted) sounds of a piano chord sequence played together (without any form of parametric alignment) with moving images metamorphosing between colours and abstract visual forms. Receivers tend to make their own subjective deductions concerning when and how the independent streams become integrated.

To summarise at this point, materials that exhibit concomitant relationships simultaneously activate (superimpose) two or more (heterogeneous) schemas. As a result of the process of integration (towards the homogeneous plane) audio and video textures become highlighted and masked. The principle of *primacy of reference* applies in these circumstances, although in situations where referential qualities are matched, dominance may be ambivalent. Where both audio and video materials exhibit referential qualities (and are heterogeneous), the presentation of audiovisual metaphors also becomes possible. The

technique of engineering metaphors has already been discussed at length in *Chapter Five: Referential Meaning in Music* (with reference to Aristotle's formula *A is to B as X is to Y*). These same methods may also be applied to audio and visual combinations of materials with equally effective results.

Isomorphic Relationships

The first species of isomorphic relationship concerns the use of congruent referential video and referential audio. A common example of this relationship is a newsreader whose expressions and lip movements precisely match the rhythm, intonation and tone of their spoken words. Another example is an image of the letter T, and the sound 'T' presented simultaneously. This type of relationship is known to be most effective as a means of delivering information. There is no misinterpretation as *the same* information is delivered simultaneously through two different (albeit integrated) media streams. Talsma, Doty and Waldorff have shown that the alignment of both form and meaning results in the fastest integration times of all audiovisual relationships.

The second type of isomorphic relationship concerns the use of congruent referential video and abstract audio. Here, the principle of 'primacy of reference' is at its most extreme, in that the abstract audio textures will normally be perceived as being entirely subservient to the referential video materials. Examples of this phenomenon may be found in early Disney cartoons, such as Steamboat Willie, or in the music of Carl Stalling who wrote extensively for Looney Tunes/Warner Brothers. In these examples, instrumental sounds were used to represent the real-world actions of cartoon characters. A violin played in a particular way, for example, might imitate speech. In these circumstances even the most remote of *abstract* audio materials may be accepted as plausible due to the intactness of simultaneously occurring formal parametric relationships.

A third kind of isomorphic relationship concerns the use of congruent abstract video and referential audio. With the definition of isomorphic in

mind (corresponding or similar in form and relations), one might imagine abstract video (e.g. particle animation) that is being *driven by* a parametrically aligned referential audio track (e.g. spoken voice). Here, where each audio onset corresponds directly with real-time video production, the principle of primacy of reference (in this case audio) comes into play as it determines both when and how the video materials will be perceived, adapting their meaning to *fit* the context provided by the audio. Fine-grained parametric alignment and multidimensional mapping procedures tend to play a major role in artificially forming convincing relationships of this type.

The final category of isomorphic relationship concerns the use of congruent abstract video and abstract audio. A feature of this category is the focus on the two-way alignment of abstract audiovisual parameters. For example, the amplitude and pitch of a stretched piano sound might be mapped to the luminosity/size and height of a *synthesised* on-screen object/image - and conversely (where the production process is centred on video materials), the luminosity/size and height of an abstract on-screen image (for example a light source) may be mapped to the amplitude and pitch of *synthesised* audio materials. The result is the production of homogenous abstract audiovisual materials that tend to be accepted as plausible due to the intactness of simultaneously occurring formal parametric relationships. While technically demanding, this range of techniques allows audiovisual artists the freedom to select a large range of abstract audio and visual materials for use in their works.

To summarise at this point, materials that exhibit isomorphic relationships activate only one schema (they are recognised as a single object or event). The principle of formal parametric alignment operates to fuse audio and video materials together as part of a natural process of integration. Furthermore, it has been established that a range of isomorphic relationships may be artificially created and controlled irrespective of abstract or referential content. This allows audiovisual artists to *impose* such relationships between remote kinds of materials, although there is always the risk that some relationships will fail. When

isomorphic relationships do fail, the result can expose the hand of the producer and highlight the relationship as a contrivance. Materials may exhibit both heterogeneous *and* homogeneous qualities in unplanned and undesirable ways. Isomorphic relationships are most effective in the fast delivery of information. As Tasma, Doty and Waldorff state, 'empirical studies have established that integrating visual and auditory stimuli serves the purpose of enhancing perceptual clarity' (Talsma, Doty and Waldorff 2006: 679). Practice-led research concerning 'perceptual enhancement' has also been carried out in several high-information-flow situations, and experimental technologies have been trialled in various commercial settings such as the New York Stock Exchange (see Pasha Roberts' *Information Visualisation for Stock Exchange Ticks: Towards a New Trading Interface*). At the heart of this research is the existence of two 'super-additive effects' that are known to occur where strong isomorphic relationships are observed. The first relates to holistic or Gestalt stimuli (increased brain activity), the second concerns fast integration time (through dual-stream delivery of information). Talsma, Doty and Waldorff explain:

> First, when the auditory, visual and audiovisual objects were attended, the P50 [an electrical wave generated by the brain] to the audiovisual stimuli was larger than the sum of the P50 activity for the auditory and visual stimuli…' [and] 'Their results, just as ours suggest that the multisensory integration effect only occurs early in time when both visual and auditory stimulus features can be constructed into a single coherent audiovisual object' (Talsma, Doty and Waldorff. 2006: 686, 688)

Perhaps the momentary acuteness of sensation we experience when attending materials that exhibit strong isomorphic relationships may be attributed to these super-additive effects.

Analysis–Reassignment Strategies

The essence of the isomorphic relationship rests on formal parametric alignment. To achieve this, in terms of a synthesis technique, requires fine-grained analysis of audio or video materials and reassignment of control data parameters to the corresponding media's parameters. Audio

data might invariably include pitch and amplitude, although third contextually relevant parameter is often required to achieve plausible results. Video data might include on-screen position, velocity of on-screen movement, screen weight, luminosity, colour, etc. Nicholas Cook adds detail to these suggested mapping strategies in stating that 'we ourselves decide which colours and sounds will best serve the given assignment or emotion as we need them… [and also that] . . . a bright colour may correspond to a bright sound in an absolute sense' (Cook 1998: 52,76). His comments encapsulate the general approach to synthesising an isomorphic relationship - that experiential schemas drawn from everyday life tend to provide the broad frameworks for effective audiovisual mapping strategies. Other researchers, such as Francisco Kröpfl, have obtained similar results in observing that the apprehension of complex systemic relationships seems to be at least partially based on the free adaptation of experiential schemas. Kröpfl reports (in relation to his experiments) the following correspondences between media types:

> A1) Dynamic correlation: envelopes related to displacement on depth and width axis of objects. A2) Degrees of brightness related with pitch registers and spectral types. B) Correspondence between structural properties of both dimensions (image and sound) such as: Regularity – irregularity, Homogeneity – heterogeneity, Continuity – discontinuity, Etc' (Kröpfl 2005: 90).

Interestingly, in dance, isomorphic relationships may be freely created by attending audio materials and 'acting out' visual counterparts. The impression is primarily one of audio-driven isomorphic integration but in rare situations (particularly with the aid of gestural sensors), this relationship may be reversed – where the movements of the dancer may be seen to control the audio materials. Leigh Landy creates this *illusion* in his work *I Conduct Electricity* (Landy 2000) by 'controlling' a series of pre-recorded sound materials (being played through stereo monitors) in various complex ways through the movements of his conductor's baton.

A related method of creating isomorphic relationships through gestural control concerns the real time synthesis of audio and video materials from a single data source. For example, the author's 2009 work *Eyepiece* makes use of an electronic device to track the unique eye movements of individual users and translate the data in real time to synthesise both sounds and moving images. Mapping strategies like those listed above are used throughout the work, and acute isomorphic relationships are realised through the real-time interactivity. During the process of developing the work, an interesting discovery was made – that the assignment of individual parameters (for example, pitch to up/down on-screen position) appears to be less important than control changes observed within the systems they constitute. Even arbitrary mapping strategies produced relatively plausible results *where parametric alignment was maintained*. The observation concurs with comments made by Kröpfl concerning the free adaptation of experiential schemas.

From a practical sound-making perspective, two further points of discussion arise. The first concerns the *complexity* of parametric alignment required to achieve plausibility (for example, amplitude only, amplitude plus frequency, etc), while the second concerns the *type* of audiovisual *mapping solution* (for example, pitch to height of on-screen image).

The first point (concerning complexity) is supported by several scholars and composers working in the related field of live and interactive electroacoustic music. The importance of presenting at least *three levels of parametric alignment* is evidenced through Traube, Depalle, and Wanderley's article *Indirect Acquisition of Instrumental Gesture Based on Signal, Physical and Perceptual Information.* They specify 'a multi-level approach for the extraction of instrumental gesture parameters taken from the characteristics of the signal captured by a microphone and based on the knowledge of physical mechanisms taking place on the instrument' (Traube, Depalle, and Wanderley, 2003: 43). In essence, the approach concerns the digital reproduction (mimesis) and extension of specific instrument types – attracting the term *hyper-instrument design.*

This audio-to-data analysis strategy may also be applied to scenarios where visual materials are *driven* by audio. Using the previous example of particle animation driven by spoken vocal materials, it is apparent that the visual materials offer an extension of schemas related to *speech.* Conversely, the technique may also be applied as a video-to-data strategy mapped to corresponding audio parameters (as the previous example of early Disney and Looney Tunes/Warner Brothers cartoons illustrate). The point to emphasise is that multiple levels of changing parameters operating within gestural timeframes are required to achieve the most convincing results.

The second point concerns the plausibility of the *type* of mapping solution that might be applied in the creation of audiovisual relationships. The audio parameter *pitch* provides an immediate example of the inherent problems in audiovisual mapping. Pitch appears to form natural isomorphic relationships with a small number of spatial paradigms including the westernised construct of up/down. Lawrence Zbikowski speaks to this topic in his 1998 paper Metaphor and Music: Reflections from Cognitive Science. He states:

> Although Scruton argued that it was virtually inconceivable to construe pitch in any way other than an up-down spatial relationship, evidence to the contrary comes from a variety of sources. Greek music theorists of antiquity spoke not of "high" and "low" but of "sharpness" and "heaviness"; in Bali and Java pitches are not "high" and "low" but "small" and "large"; and among the Suyá of the Amazon basin, pitches are not "high" and "low" but "young" and "old." The differences among these ways of characterizing musical pitch suggests that the understanding of music is profoundly metaphorical: not only is the high and low of musical pitch metaphorical, but it is only one of a number of ways to characterize pitch relations. (Zbikowski 1998: 2.3)

The human tendency to flexibly assign experiential schemas (up/down, small/large, young/old, etc.) to help *metaphorically* understand abstract musical materials gives rise to a range of creative possibilities concerning artificial mapping solutions in the fabrication of isomorphic relationships. This is an important point as it provides the basis for the creation of

several species of convincing relationships that are unlikely to occur in the natural world. It could even be said that this discovery, supplied through digital analysis-reassignment strategies provides the basis for an emergent audiovisual artform primarily based on the artificial creation of isomorphic relationships. Combining the two points discussed above, it appears that *remapping strategies* based on flexible schematic assignment (as opposed to rigid experiential frameworks) comprised of three or more dynamically assigned audiovisual parameters provide an effective means of producing and maintaining innovative and convincing isomorphic relationships.

Perceiving Isomorphic and Concomitant Relationships

The topic of this section concerns the contexts that appear to aid and disrupt the involuntary process of perceiving isomorphic and/or concomitant relationships in audiovisual materials. Four contributors have been identified. The first concerns the ability of the listener/viewer to apprehend the effects of the eight different types of audiovisual materials identified. As previously stated, audiovisual relationships tend to exhibit instantaneous combinations of concomitant and isomorphic relationships, and exactly how the combinations affect the process of integration varies greatly amongst listeners/viewers. Where one individual might perceive an isomorphic relationship that acts as a catalyst in the process of integration, another might experience the same materials in combination with a memory residue (for example, from a programme note) that for them evokes a concomitant relationship - resulting in the masking of other features. Such polar experiences are common. The second category concerns the effects of masking. Where tentative isomorphic relationships exist as part of a more complex collection of audiovisual materials, the threat of masking is ever present. In these settings it is not the absence of isomorphic relationships, but the inability to perceive them with clarity that transforms the experience from the realm of isomorphism to the domain of concomitance. When masking effects are encountered (e.g. the onset of new audiovisual materials that easily capture attention), the intactness of isomorphic relationships can be derailed. A third (related) contributing context

identified is the number of relationships (media pairs) that can be perceived simultaneously. This concept is known as bandwidth or bottlenecking. Mari Riess Jones and William Yee elaborate:

> 'Many approaches to attention build upon an information processing metaphor. An individual is viewed as a communication system through which information flows. Attention is conventionally conceived either in terms of a bottleneck which occurs at certain stages in the flow, or in terms of certain processing limits…' (Riess Jones and Yee 1993: 72).

Psychological testing suggests that humans can attend a maximum of five-to-seven simultaneous events; however, for most listeners/viewers, only three can be attended comfortably at one time. At this juncture isomorphic relationships are commonly transformed (by masking effects) and replaced by the perception of concomitant relationships. Isomorphic relationships can be restored by reducing the flow of data to below three objects/events displaying congruent and overt audiovisual parameters. Such bottlenecking principles are well known in the field of neuroscience – as Miranda explains:

> Two additional important properties to note here are: (a) it receives somatosensory inputs and (b) it has much greater processing power than lower structures, having some 400,000 neurons, as opposed to about 34,000 in the superior olivary complex (Worden 1971). Perhaps as a result of these properties, the inferior colliculus function as some sort of processing bottleneck in the auditory system. Lower inputs converge upon on it, and higher inputs to the superior colliculus (which performs early integration with the visual system), to the reticular formation (involved in autonomic processing), to the cerebellum (important centre of motor activity) and to the thalamus are projected from it. (Miranda 2010: 16)

The final point concerns the effects of interactivity. It appears that the 'primary and secondary feedback loops' (Wanderley 2000: 4) associated with *gesture* and interactivity help to maintain isomorphic relationships. For example, in the author's *Eyepiece* (2009) audio *and* visual materials are initially produced *live* by tracking the movements of the participant's eye. The audiovisual relationship is perceived as acutely isomorphic; however, in playing back the recorded materials (rather than

experiencing them in real time), the interactivity is no longer present and as such, the isomorphism is far less pronounced and therefore more open to disruption and masking effects through the presence of other audiovisual materials.

Garro's Association Strategies

It was Diego Garro who, in 2005 first proposed the term *isomorphic* to describe parametric alignment between audio and video materials. At that time, new sound and moving image editing technologies were becoming available that appeared to present potential for the emergence of new audiovisual art genres and categories. Interestingly, one such category *did* emerge, embracing the concept of isomorphism, and attracting the title *visual music* - where moving visual images were often *driven* by accompanying sounds. A significant aspect of Garro's research is the identification of several production strategies that might be represented along a continuum ranging from complete separation of audio and visual streams to total integration. One of his models is presented in figure 6.3.

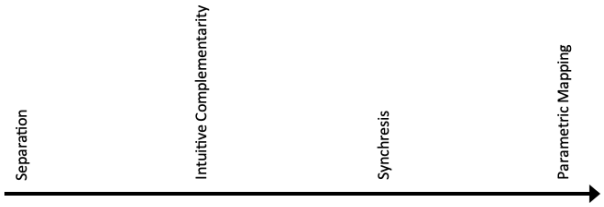


Fig. 6.3. Continuum of gestural audio vs. visual association strategies – revised version

Garro offers the following explanation of the model, which essentially presents varying degrees of audiovisual integration. He writes:

> We will also extend the continuum on the left-hand side to include cases in which there is no detectable relationship between materials in the two domains which appear to proceed on separate, unrelated phenomenological trajectories. The continuum presumes, from left to

> right, a progressively growing strength of the associative links between sounds and visuals. In most circumstances this will correspond to the experience of viewers/listeners: spectro-morphological ties in parametrically mapped passages will provide a more convincing gestural integration than synchresis. In turn, the temporal bonds in synchretic gestures will in most cases yield stronger coalescence than a loose, and often subjective, complementarity between the constituents. It should be pointed out that synchresis or parametric mapping, despite their phenomenological cohesion, might fail to accomplish gestural associations that are aesthetically plausible within the context of the work at hand. (Garro 2012: 107)

Figure 6.3 is useful in the context of the current discussion in that it can be seen to house the concepts of concomitance and isomorphism. Furthermore, It can depict other integrative effects such as synchresis and/or interactivity, and the effects of bottlenecking. Garro's chosen terms representing the two main types of integrating materials (coinciding with concomitant and isomorphic) are *allegorical* and *coalescing*. He states:

> Allegorical' strategies work on premises stronger than simple coexistence. Sounds and visuals are similarly manipulated, but not necessarily in synch, so they may be seen as metaphors of one another on the basis of general descriptive categories, even when occurring at different times within the work. (Garro 2012: 108)

And

> Coalescing' strategies are those which seek to establish audiovisual discourses principally through morphological convergence of the sonic and visual streams in most parts of a visual music work. Thus, such connection is more than merely temporal, and as a matter of fact it does not need to be, for it seeks to incorporate the two components into a recognisable pattern of phenomenological behaviours. (*ibid.*)

Basanta's Audiovisual Relations

In his 2017 *eContact* article, *Shades of Synchresis: A Proposed Framework for the Classification of Audiovisual Relations in Sound and Light Media Installations*, Adam Basanta offers constructive criticism of the model

presented in figure 6.1, together with his own improved model. In his conclusion he states:

> Having populated a three-dimensional analysis framework with eight audiovisual typologies, we arrive at a functional state-space of audiovisual relations [Fig. 6.4]. While additional typologies may be added through analysis and creative practice, the current framework significantly improves Chion and Coulter's frameworks in terms of detail. Each identified typology offers various potentials of manifestation, while remaining related to one another in terms of compositional dynamics (harmonic and horizontal axes) and perceptual meaning (axis of perceptual bond). Throughout the identification and placement of typologies within the three-dimensional framework, I have maintained a focus on the relation between compositional parameters and perceptual attributes. This focus allows an examination of the interaction between the two, especially in terms of the various strengths and "shades" of synchretic experience. (Basanta 2017: 23)

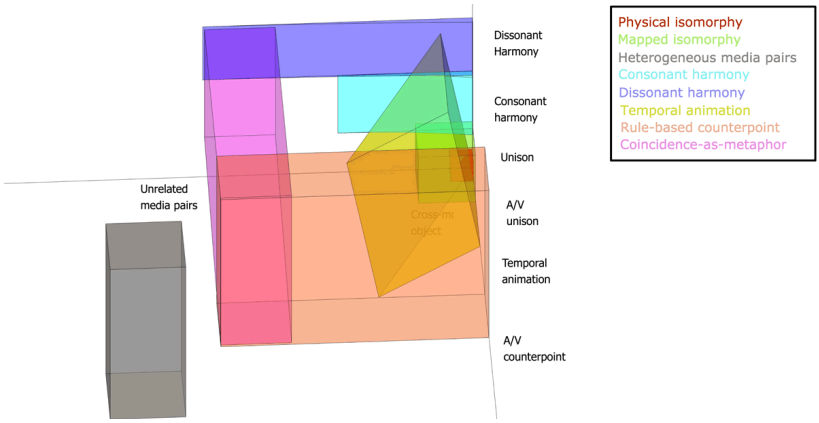


Fig. 6.4: Basanta's Audiovisual Relations Model (ibid.)
(reprinted with permission from the author)

Basanta's model (presented in figure 6.4) puts forward several advancements on the figure 6.1 model discussed above including a splitting of the term isomorphic into *physical isomorphy* and *mapped*

isomorphy, which directly supports the assertions made above regarding *nature as the architect* of such relationships. The red box (physical isomorphy) depicts the most natural type of isomorphic relationship, while the green box (mapped isomorphy) extends the relationship type through remapping strategies resulting in the creation of plausible audiovisual constructs that are unlikely to exist in the natural world. At the other end of the continuum, and of particular interest, is the way in which Basanta separates-out heterogeneous audiovisual materials (the grey box to the left). He states:

> The polar opposite of the homogenous, isomorphic typologies can be defined as heterogeneous media pairs. In this typology, audiovisual relations are perceived to be different in both "harmonic" content and temporal development, resulting in a lack of cross-modal binding. This is not to say that heterogeneous media pairs do not share mutual influence, but rather, that this mutual influence is a type of "added value" rather than a synchretic experience (Chion 1994, 5) (Basanta 2017: 12)

This observation mirrors both my own (2010) and Garro's (2005, 2012) acknowledging that in rare situations audio and visual attention may be divided. Returning to the paradigm of isomorphism, interestingly, Basanta's concept of consonant and dissonant harmony appears to rely on the maintenance of parametric alignment (onset, termination, morphology, etc.) to create a continuum between plausible and contrived relationships. Regarding consonant harmony he states:

> We are aware that the combination of light and sound is contrived; however, we allow ourselves to indulge in the illusion of unity, as the temporal unison of light and sound media make the audiovisual events ecologically plausible. (Basanta 2017: 14)

Regarding dissonant harmony he states:

> Despite [the] onset synchrony and morphological symmetry between light and sound intensity, the lack of appropriateness between sound and light events challenges the perception of a bonded cross-modal object due to the concomitant relationship between media pairs. Using Coulter's terminology, the two schemas triggered by this event (a light bulb and a

> piano [the given example]) are so different in terms of content that highlighting may only occur on the perceptual level, not on the conceptual level. (Basanta 2017: 15)

Here, Basanta splits the duality described in the definition of *isomorphic,* as 'similar in *form* (parametric alignment) and relations (referential associations) maintaining formal parametric alignment and allowing associative *contrivances* to be deliberately presented and exposed, which in turn insists on some form of reconciliation. The concept of formal *mimesis* is at the heart of the technique. Regarding the temporal axis, Basanta recognises the importance of onsets in media pairs, and emphasises appropriate levels of departure in terminations. Furthermore, he demonstrates how synchresis can play a part in creating the impression of temporal alignment where none is present (parametrically). The category of audiovisual counterpoint is also an interesting one, as it relies on the establishment of *rules* (such as repeated patterns) that may be freely shared, inverted or adapted so long as the basis of the rule is maintained. Basanta's final category of *coincidence as metaphor* is closely aligned with the concept of concomitant relationships as discussed above. It addresses 'light and sound behaviour [that] coincide[s] in time but do[es] not exhibit morphological congruency' (ibid). The coming together of audio and visual streams into one coherent audiovisual stream again appears to rely on a less pronounced form of Chion's synchresis.

A New Multimodal Model

Within the narrow context of digital audiovisual production, a feature of all three models examined (Coulter's, Garro's, and Basanta's) is the way in which they describe a range of audiovisual relationships ranging from disassociated, to fully integrated. They all tend towards a point of *ideal combination* resulting in a single *audiovisual* stream (presenting one schematic association). The models describe the coming together of two disparate types of media (audio and visual) to create a third type (audiovisual) featuring Gestalt qualities. It is also recognised that *synchresis* plays a part in this process (particularly in non-ideal combinations), resulting in simultaneously attended *integrated*

(highlighted) and *non-integrated* (masked) component materials. This fundamental process depicting audiovisual integration is represented in its most basic form in figure 6.5.

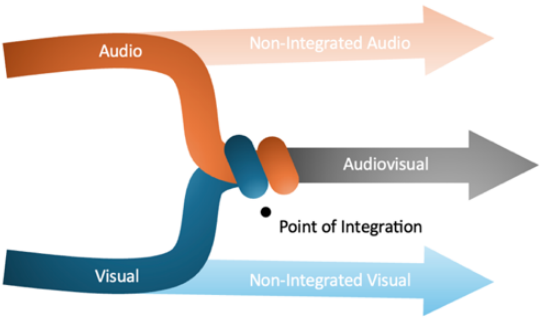


Fig. 6.5: Audiovisual Integration

The next level of detail concerns the way in which certain combinations of media come together. All three models specify a range of relationships (presented through the materials themselves) that act out in a variety of ways in the process of forming integrated audiovisual materials (for example, the Necker cube identifies eight such relationships). Furthermore, all three models offer a similar *ranking* in terms of the integration capacity of certain types of relationships. *Isomorphism,* for example, (a term used by all three researchers) is collectively regarded as a powerful integrating principle that may be used in the service of several techniques that create the impression of a single object or event (the activation of one dominant schema). By comparison, methods of complementarity (Garro), concomitance (Coulter), or heterogeneity (Basanta) are regarded by all three researchers as less integrated due to the way in which they present two schematic associations (creating the impression of two superimposed objects or events) Here, *common characteristics* tend to be highlighted in the presentation of audiovisual metaphors – a process somewhat reliant on *synchresis*.

In terms of graphically representing the integrating qualities that may be found within audio and visual materials, I concur with Basanta (see figure 6.4), who artificially assigns *values* to several *axes* representing perceived similarities and differences – that is, where no differences are observed, ideal integration approaches zero (*physical isomorphy*). Figure 6.6 presents a similar approach with reference to each of the eight entangled modes of listening identified, and the roles they play in audiovisual integration.

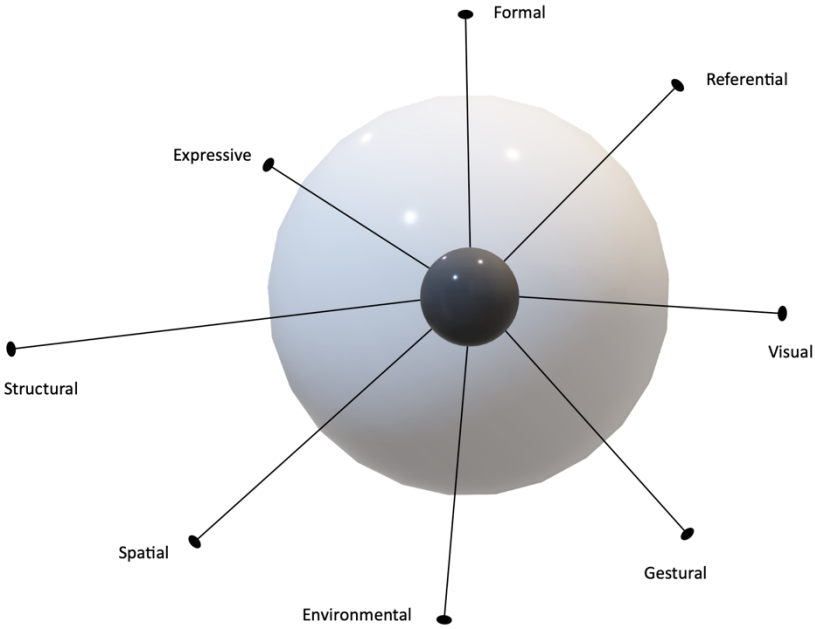


Fig. 6.6: A Multimodal Network of Audiovisual Integration Fields

Homogeneity (isomorphic relationships) might be considered naturally commensurate and therefore might be positioned at the centre of the model, with heterogeneity (concomitant relationships) naturally incommensurate, represented on the periphery of the model. Using the eight modes of listening/seeing identified (expressive, formal, referential,

visual, gestural, environmental, spatial, and structural), we might imagine numerous intersecting axes (spherical harmonics) coinciding at the zero-point. Similarities and differences between audio and video materials might then be considered in multiple dimensions. For example, referential qualities might remain different, where formal isomorphic qualities might be fully integrated (as in Basanta's example of the light bulb-piano). The key point to consider when undertaking any such analyses on any axis is the *level of difference between the two contributing media types.* The greater the difference, the further from the zero-point on that particular axis. For example, the Letter 'T' and the sound 'T' are both highly referential (delivering the same information), and so on the referential axis would be measured at a point close to zero (the centre of the model). The paradigm might be further conceptualised as a series of *tests* measuring both audio and visual qualities with respect to each of the identified modes of listening. This differs somewhat from the more general approach taken in chapters 3-5 and 7-10, which seeks to expose *multimodal blends* through the eight entangled modes of listening identified. Detail concerning each *test* (with respect to each of the modes of listening identified) will now be supplied.

The Expressive Mode of Listening/Seeing

This field concerns basic emotional likes and dislikes. We have already discussed the primacy of expressive meaning in music and in audiovisual contexts a similar scenario is apparent. Emotive linking qualities such as *sad* (shared by image and sound) tend to present themselves relatively early in the listening/seeing process in comparison to linking qualities that might arrive through more sustained modes of auditory and visual attention. In film contexts this type of *auditory* functionality is often referred to as *mood.* In the context of the current discussion, *expressive* similarities and differences between constituting audio and visual streams might be contemplated through the feelings they evoke.

The Formal Mode of Listening/Seeing

The term *formal* may also be expressed as *morphological* or *abstract,* highlighting onsets, terminations and other shared formal qualities such

as intensity envelope, size, shape, position, etc. This mode of listening/seeing, relevant to both audio and visual materials, essentially concerns the recognition of shared *patterns.* Interestingly, in audiovisual contexts, shared formal qualities tend to dominate the process of integration, and in abstract contexts seem to be largely responsible for the formation of isomorphic relationships. Parametrically aligned attributes such as these are heavily integrative, even where other accompanying meaning-based qualities remain non-integrated. Formal qualities may also be appreciated for their musical or artistic pattern-bearing merits in isolation from their capacity to integrate audiovisual materials.

The Referential Mode of Listening/Seeing

The term *referential* refers to the *recognition* of an object or event through the associative meaning-bearing qualities of both sounds and images. For example, the image and sound of a car both carry the referential meaning '*car'.* The term may also be considered with reference to various levels of departure beginning with an intact *source-bonded* association, progressing to a range of *source-broken* associations, and ending with a wide range of weak associative possibilities. Smalley has referred to this gradual degradation in object/event recognition as 'surrogacy' (Smalley 1997) – progressing through 1st order, to 2nd order, to remote surrogacy. Where no common object or event is identified, the audiovisual construct may also present dual referential meanings through various levels of *divided attention*. The category highlights the '*what'* characteristics of both sounds and images.

The Visual Mode of Listening/Seeing

This multisensory mode of listening/seeing refers to the visual sensory experience (either imagined or real) that results from the reception of auditory and visual information. The test may be applied to all audiovisual contexts, including concertised musical performances, where, like mediated experiences, conflict can arise between quasi-visual associations supplied through audio materials, and actual visual experiences. This scenario commonly arises in the presentation of

acousmatic music, where visual materials of any kind are often avoided due to the likelihood that they will be incommensurate with the complex *quasi-visual soundscapes* supplied through the sound-world. Here, dramatic differences between audio and visual experiences are self-evident, and in contrast with earlier examples of integrated audiovisual materials where mediated visual images tend to *match* quasi-visual images produced through associated audio streams.

The Gestural Mode of Listening/Seeing

This multisensory mode of listening/seeing supplies powerful associations with *performative agents* – that is, sounds and/or images that are perceived as created by a human protagonist. For example, an on-screen dot, whose x/y/z position is produced by gestural hand movements (or in audio terms sounds of the human voice) tends to automatically result in anthropomorphism of visual materials. This species of integrative relationship may be experienced either in mediated form, or directly (in interactive installation or virtual reality contexts where the gestural movements of participants are tracked and converted to sounds and/or images in real time). In a wider sense, the category also embraces impressions that multiple *agents* may be responsible for the genesis of various collections of sounds and images as well as the various types of behaviours that might be observed among them. The current context requires the strength of *gestural associations* to be identified in both audio and visual streams.

The Environmental Mode of Listening/Seeing

An environment is a construct that has spatial dimensions and associations with specific places drawn from life experiences. Our evolutionary need to make sense of our environment as a unified spatial construct containing identifiable objects or events may very well be at the heart of our natural tendency to integrate audio and visual information streams. The integrating power of this field means that audio and visual spaces do not necessarily need to be matched, as suitable schemas will tend to be adopted to reconcile differences (for example, foreground/background). The environmental field also tends to be

heavily entangled with other fields including spatial, visual and referential modes of listening/seeing; however, for the purpose of this discussion may be conceptualised as the level of environmental association brought to bear through constituting audio and visual streams.

The Spatial Mode of Listening/Seeing

Spatial intactness and/or separation becomes a directly tangible concept when considering the practicalities of audiovisual projection. This is especially the case when bearing in mind the three-dimensional nature of sound in combination with two-dimensional on-screen images. Other abstract reductions of space in the form of image schemas have already been discussed, and within the current context these may also be viewed as integrating concepts relevant to artmaking strategies. The up/down image schema's association with high/low sounds, for example, may be directly applied within the context of managing high/low on-screen images. More complex spatial associations also tend to operate with the aid of the time domain. These include the concepts of *centrality* and *pathways and goals.* The similarity/different test may be carried out through considering the range of spatial impressions that arrive through both audio and visual materials.

The Structural Mode of Listening/Seeing

The structural axis addresses the way in which audio and visual materials are perceived in terms of their *organisation*. Operating over longer timeframes, the structural field in the current context primarily relates to patterns perceived over the duration of entire works. For example, the way in which an opening audiovisual sequence might be repeated towards the end of a film or multimedia artwork. Interestingly, a replaying of audiovisual materials that has already been subjected to a process of integration, encourages a replaying of the same listening/seeing strategy that integrated them in the first place. The identification of structural similarities/differences between audio and visual streams also gives rise to a range of production techniques relevant to *closure.*

The Multimodal Network Model and the Five Layers of Mental Activity

In the process of cross-referencing the multimodal network model presented in figure 6.6 with the five layers of mental activity: sensations, gist, locus, contexts, domains, the concept of simultaneous activation and entanglement becomes evident. While formal isomorphism (attended at the level of sensations or gist) might provide the impression of intactness (homogeneity), a competing field, such as the referential qualities might yield a difference (heterogeneity) and interestingly, these two impressions may be simultaneously activated, yet remain unreconciled. Such impressions are common within the context of experiencing audiovisual art, and this can insist in the plausible creation of novel and remote relationships (as Basanta attests to). In such scenarios, the disparity between integrated and non-integrated audiovisual components does not necessarily register an error, as these types of understandings are likely to be processed by different parts of the brain corresponding to the five layers of mental activity previously noted. It is not until these centres are more thoroughly examined that discrepancies become apparent. For a more detailed explanation of the five layers of mental activity and the way in which they interact please refer to *Chapter Two: Multimodal Listening*, figure 2.2.

Concluding Comments

Let us return to the paradox presented at the beginning of the article – that audiovisual materials on one hand offer the possibility of complementing one another, and on the other, carry the risk of detracting from one another. It has now been determined that these conflicting statements are simple descriptions of homogeneity and heterogeneity. In the first part of the chapter, the author's 2010 model is put forward offering a classification system for combinations of abstract and referential audiovisual relationships. Eight examples are provided to illustrate the functional qualities of each category (one from each corner of the Necker cube). Two main types of relationships are identified: 'concomitant' and 'isomorphic'. Concomitant relationships rely on the highlighting and masking that occurs when two schemas are simultaneously activated (superimposed), while isomorphic relationships

defined as *similar in form and relations* result in the activation of a single common schema. The two relationships may also be respectively described as heterogenous and homogenous. Detailed methods of creating both concomitant and isomorphic relationships have been proposed. Four agents that aid and disrupt the perception of isomorphic and/or concomitant relationships have also been identified. These include: the attention of the listener/viewer, the effects of masking, of bandwidth or bottlenecking, and of performative interactivity. The model is then cross referenced with other models drawn from the field of visual music including the work of Diego Garro (2005, 2012) and Adam Basanta (2017), who add concepts such as a *continuum of association strategies* (Garro), *physical* and *mapped isomorphy*, and *consonant and dissonant audiovisual harmony* (Basanta). The research of Garro and Basanta also provides the groundwork for comparison and improvement to the initial model. Of particular interest is the proven existence of *super-additive effects* associated with audiovisual integration (Talsma, Doty and Waldorff 2006) – a phenomenon that Garro also attests to in the words of Pellegrino. He states:

> The principle of synergy, i.e. that the whole is greater than the sum of its parts, is fundamental to the nature of complex interactive dynamic sound and light systems and their resulting forms (Pellegrino 1983: 208) - quoted by Garro 2012: 108).

In the second part of the chapter, a new multimodal model is proposed incorporating eight modes of listening (and seeing). The eight fields (represented as spherical harmonics) include: expressive, formal, referential, visual, gestural, environmental, spatial, and structural qualities. Each mode assists in measuring similarities and differences in audio and visual materials. The model is then cross referenced to the five mental layers of activity model supplied by Kendall (2014), which highlights the likelihood of all integrating modes operating in parallel, managed by multiple areas within the brain that are simultaneously active. The construct explains how one species of audiovisual relationship may be appreciated as fully integrated (for example, formal isomorphism), while another (for example referential) remains

incongruent. It seems that such divided impressions of this type are extremely common within the context of attending fabricated audiovisual materials (less so in natural settings). There also appears to be a natural *ranking* of the eight identified fields with formal pattern recognition and parametric reassignment taking a leading role in the process of integration (naturally forming isomorphic relationships). Concomitant processes (where differences are observed) take less of a role in integration but nevertheless enable a range of powerful techniques such as the presentation of audiovisual metaphors and other combinations reliant on the principle of *synchresis*.

The narrow outcomes of the chapter offer an advanced multimodal network model, supplying a vocabulary, and classification system relevant to the *fabrication* of numerous types of audiovisual relationships within digital media production contexts. The investigation also reveals the transformational power of visual images to support, highlight, mask, or unhinge communicative sonic characteristics bringing to light important differences between acousmatic (audio-only) and audiovisual modes of attention. These wider findings extend to all other sound-based domains that incorporate visual experiences including acousmatic music and live performance, providing new understandings of these art forms and their communicative potentials.

This shift in thinking brings with it the necessity of considering the impact of visual images on music and sound in a variety of settings including those traditionally regarded as audio-only artforms. The principles investigated in this chapter (through the narrow lens of audiovisual media pairing), indicate that it is likely that the relationships identified also form the basis of quasi-visual image creation in audio-only settings. That is, the conjuring of imaginary images to accompany sounds is most likely to be informed by *affordances* presented by the eight entangled modes of listening identified - the expressive, formal, referential, visual, gestural, environmental, spatial, and structural qualities of sound. The special focus on imaginary images may also be described as *visual blends* (between the visual, and all other modes of listening). In contemplating

the wide range of quasi-visual images that may be suggested through these associations, a staggering variety of hypothetical possibilities becomes apparent. Furthermore, one of the more surprising discoveries of the chapter is that the combination of sounds and images (real or imagined) need not be strictly congruous, as different layers of mental activity and separate areas within the brain are responsible for managing the constituting materials. It seems that strong audiovisual relationships may be formed even where one or more of the identified modes of listening/seeing is *incongruous*. The chapter has revealed that the *visual mode of listening* is far more active, displaying a greater level of entanglement and variety of accompanying images than previously imagined, with the possibility of up to several congruous/incongruous real and/or quasi-visual images contributing to the listening experience at any given time.

Chapter Seven: Gestural Associations

This chapter addresses the topic of musical appreciation through association with the human body. The concept of *gestural agency* will first be discussed in the widest context, as an entangled component of the listening experience applicable to all genres and categories of music. The chapter also has narrow implications, relating to a subset of sonic art incorporating both live acoustic and electroacoustic sources. These practices go by many names including *mixed works* (referring to a *mix* of instrumental and electroacoustic music) and *live electronics* (instrumental music with live electronic processing). In the second part of the chapter, detail will be supplied concerning the *reconstruction* of plausible gestural relationships within these *live sonic art* settings. These wide and narrow perspectives on *gestural agency* combine to provide a detailed overview and understanding of the paradigm.

Clarification of the term *agency* is first offered. In *Chapter Three: Musical Expression* it was defined as referring directly to the *agent* (normally human) who is perceived as being responsible for creating the musical materials (note, sound, phrase, event). In essence, the paradigm relates to the conservation of energy (as a transformation pathway), normally beginning with kinetic (human gestural, environmental, mechanical, etc.) or electrical energy and ending in the production of sound. This method of reconciling musical impetus is extremely common irrespective of musical style or genre – as Gary Kendall explains:

> In the everyday world, listeners identify agency with the cause of the changes of state associated with the event. We can readily understand that most types of Events have a slot for agency. In the case that the agency of an event is unknown or unknowable, there are potential default values. In live performance, the default agent is likely the to be the performer. For many listeners of electroacoustic concert music, the default agent is the composer. Interestingly, many listeners associate electroacoustic music with physical gestures and human agency even when no physical activity was involved in its creation. This is in some part a reflection of the

> importance of agency to the everyday mind's sense of meaning. (Kendall 2010a: 67)

Gesture is a concept that operates in parallel and is entangled with the concept of *agency*. As previously noted, it may be appreciated in *surrogate* form and constitutes an archetypal component of the listening experience. The *gestural mode of listening* proposed and discussed in this chapter absorbs Smalley's (1996) *gesture, utterance*, and *behaviour* fields (although it is limited to the human elements of his behaviour field). It is used as a broad catch-all for all associations with human agency. The term *Gesture* refers to the energy pathway itself (the perceived *envelope* or ASR curve), whereas *agency* might be assigned to the protagonist (real or imagined) who produces the gesture. Denis Smalley provides the following useful description of *gesture*, which also specifies a range of *multisensory* qualities and associations:

> A gesture is therefore an *energy–motion trajectory* which excites the sounding body, creating spectro-morphological life. From the viewpoint of both agent and watching listener, the musical gesture-process is tactile and visual as well as aural. Moreover, it is proprioceptive: that is, it is concerned with the tension and relaxation of muscles, with effort and resistance. In this way sound-making is linked to more comprehensive sensorimotor and psychological experience. … Not only do we listen to the music, but we also decode the human activity behind the spectromorphologies through which we automatically gain a wealth of psycho-physical information. (Smalley 1997: 111)

Smalley's *energy and motion trajectories*, discussed in chapter five in terms of 'converging/diverging, ascending/descending contours, and accumulative/dissipative and diffractive/conglomerative growth' (Smalley 1996: 88) bear a striking similarity to Snyder's (2000) image schemas: up and down, centrality, motion- linkage-causation, linearity: pathways and goals, and containment: inside and outside, and this is hardly surprising, given that image schemas are also derived from associations with the human body. *Gesture* then, and its inherent relationship with image schemas may be seen to occupy one of the most

basic ways in which we as humans have come to appreciate music, and wider still, the world around us – as Lakoff and Johnson explain:

> Spatial-relations concepts are embodied in various ways. Bodily projections are obviously based on the human body. Concepts like front and back … arise from the body, depend on the body, and would not exist if we did not have the kinds of bodies we have. The same is true of fundamental force dynamic schemas: pushing, pulling, propelling, supporting, and balance. We comprehend these through the use of our body parts and our ability to move them, especially our arms, hands, and legs. Other image schemas are also comprehended through the body. Our bodies are containers that take in air and nutrients and emit wastes. We constantly orient our bodies with respect to containers [such as] rooms, beds, buildings. We spend an inordinate amount of time putting things in and taking things out of containers. We also project abstract containers onto areas in space, as when we understand a swarm of bees as being in the garden. Similarly, every time we see something move, or move ourselves, we comprehend that movement in terms of a source-path-goal schema and reason accordingly. These forms of embodiment arise from the way we schematize our own bodies and things we interact with daily (Lakoff and Johnson 1999: 43)

Within musical contexts, the fact that both musicians and audience members react involuntarily to shared patterns such as *tension and release* provides powerful evidence of the psycho-physical effects of music and its direct relationship to the human body. This proprioceptive relationship is thought to form a large part of abstract musical functionality, resulting in *physical* (as well as psychological) reactions to common compositional techniques such as *harmonic cadence*. Gary Kendall provides further comment:

> In everyday life, the experience of events is interwoven with the flow of felt experience. Similarly, Johnson, speaking of the embodied nature of image schemas, explains that bodily movements are characterised not only by their image schemas, but also by their 'distinctive qualities' (Johnson 2007: 21). These he relates to the four dimensions of body movement described by Sheets-Johnstone (1999): tension, linearity, amplitude and projection. This terminology could equally well be applied to the flow of energy in musical events. … The embodied basis of such energy flow explains how

> listeners so easily relate sonic events to physical gestures. The felt qualities of experience are clearly recognised by listeners, but words often fail to capture the nuances and immediacy of direct experience. (Kendall 2010a: 66)

From a narrow perspective, *gesture* relates explicitly to the communicative language associated with music and sound in live performance settings. The scope of the visualised gesture to sounds paradigm encompasses all genres of performing arts including music, dance, theatre, opera, and circus arts; however, much detail may be gleaned from focusing on the even-narrower context of instrumental performance with live electronics. Beyond the realm of acoustic sound production (where energy is automatically conserved) gesture-to-sounding-result relationships must be recreated, and to achieve convincing outcomes, an advanced understanding of constituent relationships is required. In these settings, acoustic, and mediated sound, together with visual experiences combine to create *local* and *field* relationships (Emmerson 2007) that underpin artistic communicability.

This narrow paradigm will be explored with the aid of Emmerson's elaborated model (presented in figure 7.1), together with several other advanced approaches to digital musical instrument design and interactive music making. Numerous examples of creative work will be used to illustrate the concepts put forward. Although much of the detail presented in the chapter flows from a deep investigation of acoustic instrumental performance with live electronics, many of the discoveries made are also highly relevant to the more general topic of *gestural association* in a wide variety of artmaking practices including all sound-related genres and categories.

Background

In 1998 in *Composition, Performance, Reception* Clarke and Davidson put forward the idea that human gesture (centred on the body) is of utmost importance in the communicability of live instrumental music. From their studies they were able to deduce that live instrumental music is produced

through a single mechanism they termed the 'motor programme to music continuum'. They state:

> Davidson (1993, 1995) has demonstrated that information about both structural features and expressive intentions is communicated to observers through body movement. "… when both aural and visual information were used as stimuli, non-musicians relied almost entirely on the visual information from the body movements for their judgments of the performer's intentions. (Clarke and Davidson 1998: 74)

and

> The body is not just a source of sensory input and a mechanism for effecting output. A wider range of other factors, including the possibilities of the instrument, the acoustic of the performing environment, the nature of the audience, the mood and intentions of the performer, and even the performance ideology espoused (e.g. 'historically aware' performance practice) will contribute to the result, sometimes to the detriment of structure. Movement, and the human body, are particularly significant in this complex set of relationships for all sorts of reasons – the most obvious being that music is produced by human and instrumental movement and is thus indelibly stamped with its bodily and instrumental origins. (ibid)

In the year 2000, in an interview with Larry Austin and Denis Smalley, Smalley argued that in the context of 'mixed works' (for live instrument and tape), the language central to communicability is *audiovisual*. He states 'the performer is the centre of focus – and the sounds the instrument makes'; that 'the performer cannot be subsidiary to the electronics'; and that different rules for different languages (audiovisual vs acousmatic) apply – raising several questions around how we as composers might create coexistent scenarios. From these early studies we are provided with the beginnings of a functional model for producing effective *mixed works* for the concert hall. Summarising Smalley's main points, the language of live musical performance is *audiovisual* - centred on gesture, on the body, on the performer as the focus of attention, and music somehow spreads out from this central position. This deduction has significant implications including the need to manage audiovisual relationships in the ways indicated in chapter six.

Emmerson's Local/Field Relationship

In his book *Living Electronic Music* (2007), Simon Emmerson proposed an important model that sought to encapsulate the language of live performance *with* electronics. He began by examining the human body as the evolutionary *source* of music, noting heartbeats, breath, and movements of limbs as the frameworks for gestural timeframes and the patterns and cycles of nature as environmental timeframes. This juxtaposition of timeframes and associated agents led to the proposal of his *Local/Field Relationship* where:

- Local controls and functions seek to extend (but not break) the perceived relation of human performer action to sounding result
- Field functions create a context, a landscape or an environment within which local activity may be found.

Local space may also be described as *nested* within field space. Emmerson's model (together with elaborations from the author) is presented in figure 7.1.

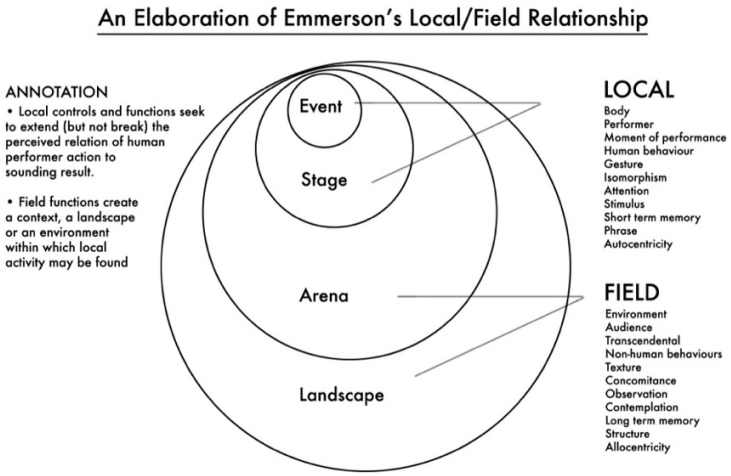


Fig. 7.1 An Elaboration of Emmerson's Local/Field Relationship (Emmerson 2007)

The Local Relationship

At the heart of Emmerson's *local relationship* is the perception of a cause-and-effect relationship formed by the performer's gestures (often attended visually) to the sounding result. This may also be described as an isomorphic audiovisual relationship (see chapter six). As previously mentioned, while such relationships often occur spontaneously as part of the production of acoustic music, in mixed settings (involving both acoustically produced materials and mediated materials) they must be *fabricated.* They do not occur spontaneously as they do in natural settings. This is a highly significant deduction for all creatives working with music technology as it means that to present a convincing local relationship, artmakers must construct plausible cause and effect relationships between the performer's gestures and mediated sounding results. Such models are invariably based on acoustic instrument designs – as it is those physical sound-makers that are naturally subject to the laws of conservation of energy. This in turn provides audiences with a plausible source of agency, marking the beginning of developmental pathways that assist in proceeding from local to field relationships.

The Field Relationship

As perception moves from the event to the stage to the arena to the landscape, musical gestures tend to become more distanced and disconnected from their bodily origins. This requires an alternative method of reconciling the cause-effect relationship. As such, *field agents* are often perceived as originating from spaces that are not occupied by performers. They come from *elsewhere*. This phenomenon has already been discussed at length in *Chapter Three: Musical Expression,* but given the audiovisual context, here we might also emphasise the emergence of *concomitant relationships* (see chapter six) where more than one *cause* or agent may be observed. As the performer's actions and movements become non-commensurate (heterogeneous) with the sounding result, cause-effect relationships can no longer be maintained, and alternative *causes* (sources of energy) are naturally sought out by listeners. These *field* sounds and their corresponding *agents* are ordinarily perceived as accompanying the *local* musical performance.

Gestural Agency

As discussed, in performing arts settings the *local* relationship automatically produces what we have termed *gestural agents*. The music is normally perceived as being produced by a *real* performer (or performers) through their on-stage presence. A *field* relationship presents a landscape from within which local activity may be found. Field relationships also contain *imaginary* gestural agents that can range from human agents of all types to mechanical and electrical energies, to environmental energies (such as the elements), to the sustaining qualities of large empty spaces (reverberation). In *mixed* contexts (employing both acoustic instruments and mediated sounds) the deliberate presentation of these *other* agents is ubiquitous. They are the multitude of associations that accompany a live performance – often revealing a relationship between the performer and their environment. In his *Organised Sound* article 'From Sound Shapes to Space-Form: investigating the relationships between Smalley's writings and works' David Hirst discusses the primacy of gesture with reference to Smalley's concepts. He writes:

> Smalley warns that if one strays too far from these cultural archetypes of gesture, the music will appear to be 'cold, difficult, even sterile' (Smalley 1997: 112). It is interesting that Smalley relies on the 'detection of humanity' to define gesture, and in particular by reference to instrumental and vocal music archetypes. Much electroacoustic music includes gestures derived from the natural world, including those gestures with animal origins, and those with links to the forces of nature such as the weather (wind, rain, thunderstorms). There are also those gestures that derive from the 'modern everyday world'. Some may have human gestural associations (e.g. work activities) or gestures associated with the built environment (e.g. machine gesture patterns). Smalley does not include these in his model. (Hirst 2011, 44)

Hirst's comments highlight what Smalley considers to be at the heart of *gesture* – association with the human body. The passage presents an interesting question – whether *all* sounding gestures (the sound of an accelerating car for example) might be perceived in *human* terms (anthropomorphised), or whether several categories might make themselves apparent (human, environmental, electrical, mechanical,

etc.). However, given evidence that has recently come to light (such as the five layers of mental activity model and the eight entangled modes of listening identified) this type of compartmentalised thinking may be irrelevant. It can be assumed that expressive, formal, referential, visual, environmental, spatial, and structural modes of listening will be naturally *entangled* with gestural associations, and that important aspects of all modes may be highlighted through the lens of *gestural agency*. What listeners receive (in general), will more likely be a blend of all *modes* and *layers* driven by their own subjective life experiences, which of course will include *gestural* understandings of the world.

The middle-ground between the local and the field is a potent one, as it naturally produces ambiguities between the two types of performative agents (local and field). The complete spatial construct (event, stage, arena, and landscape) is one where (in performing arts settings) real performers are almost always in residence, but where imaginary agents may come and go - where relationships between the local and the field spontaneously form and collapse. At the heart of this binary division between the local and the field is the perception of homogeneity *or* heterogeneity. In audiovisual terms, such relationships have already been described as *isomorphic* and *concomitant* drawing attention to issues of parametric alignment between vision and sounding result as the cause of the distinction.

Wanderley's Digital Musical Instrument Model

Marcello Wanderley supplies detail that is useful in the realisation of isomorphic and/or concomitant relationships within the context of live performance. In an article entitled *Gestural Control of Music* (2000) he suggests that 'primary and secondary feedback' should ideally be equal, and gesture (measured in terms of total perceived energy from the performer) must be reconciled with sounding result. His model is presented in figure 7.2.

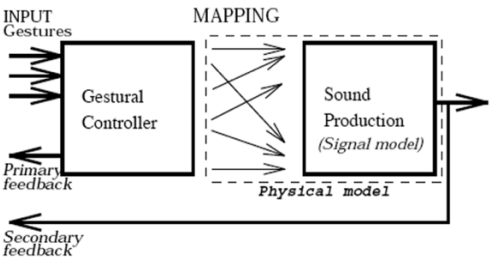


Fig. 7.2 A Digital Musical Instrument Representation (Wanderley 2000)

There is much that can be gleaned from this graphical representation with reference to the *physical model* on which Wanderley proposes a digital musical instrument should ideally be designed. The approach embraces multiple measured input gestures, and the laws associated with conservation of energy – where greater/lesser input gestures automatically result in greater/lesser sounding outputs (perhaps in terms of volume, pitch, timbre, etc.) Here, spatially derived *image schemas* (as discussed in *Chapter Four: Formal Musical Qualities*) might also be employed to help guide novel creative ideas. Of particular interest is the crossing of arrows that suggests the *reassignment* of some (but not all) of the gestural and sonic parameters found within these naturally occurring schemas together with the conservation of control data (as opposed to trigger data) in terms of phraseology of gesture to dynamically changing sounding result. Another way to describe this ideal relationship is that the ASR curve of the input gesture should match the ASR curve of the sounding result. The concept of transference or conversion of visually attended gestures to sounding result is an important one as it is likely that receivers *follow* such pathways to establish isomorphic relationships – a Suk-Jun Kim explains:

> But what kind of sounds in electroacoustic music, other than those that emanate directly from a human body, can suggest to listeners a sense of human action? Categorising sound-related gestures – sound-producing gestures, sound-accompanying gestures and … affective or emotive

> gestures – Godøy notes that the most 'embodied' gestures among these categories are 'those that follow the sound closely, i.e., the sound-producing and sound-tracing gestures', and also maintains that '[the] sound-producing gestures have an energy transfer from the performer to the instrument, whereas the sound-tracing gestures may mimic excitatory gestures as well as trace the evolution of the resonance of sounds, i.e., the ''passive'' or energy-dissipating phase of the sounds, hence not transferring energy to a resonating body' (Godøy 2006: 154).(Kim 2008: 49)

Paine's 'Category Mistake'

Garth Paine adds fine-grained detail to Wanderley's model in proposing three models of interactivity: conductor, chamber music, and improvisational (ideally conversational). He also claims that mapping human gesture to musical sounds (the pitch-time lattice) is a *category mistake* – implying that gesture to timbre (sound-based music) is far more satisfactory. He states:

> I suggest that the imposition of musical characteristics on interactive music systems only makes sense when designing a system as an accompaniment or improvisatory application where the other elements of the ensemble are adhering to chromatic music performance practice. Accepting a pre-existing musical framework as the criterion for the design of an entirely new musical paradigm clearly restricts the selection of development criteria. If one accepts that the aim is to create new means for musical expression, to generate new or innovative music, then the acceptance of historical criteria associated with chromatic music (i.e. pitch, pitch structures that define tonalities, the recognition of chords, the recognition of tempi and time signatures, etc., as out- lined by Rowe (1993) as the characteristics of an interactive music system) is inappropriate. The field of philosophy defines this kind of criteria confusion as a Category Mistake (Ryle 1995), i.e. the application of characteristics, descriptive of one phenomenon onto another where they are not necessarily characteristic. Perhaps it is time to develop a new conceptual frame- work for interactive music systems (Paine 2003: 300).

If we are to apply Paine's advice literally, then we might produce mapping strategies between a range of human gestures and previously discussed *sound-based* characteristics (as opposed to lattice-based characteristics) that constitute music and sound. Although it is not explicitly stated, here

Paine was most likely referencing Smalley's (1996) nine indicative fields with a focus on timbre and spectromorphology. His approach expands significantly on previous gesture-to-sounding-result models that aim to emulate acoustic instrument designs. Such strategies might be contemplated with the aid of the following table.

Gesture (of performer)	Smalley's (1996) Nine Indicative Fields (sound-based sounding result)
• Breath (duration, pressure) • Heartbeat (rate, fluctuation) • Pre, present, and post gestural envelopes (ASR curves) • Movement of limbs (x, y, z position, rate of movement) • Onstage location (x, y, z position, rate of movement)	• Gesture • Utterance • Behaviour • Object/Substance • Energy • Motion • Environment • Vision • Space

Table 7.1: Gesture to Sound-Based Sounding Result Strategies

For Paine, the term *agent* applies not only to the performer or person interacting with an installation, but to the computer as well (energy from the computer). He indicates that the emergence of such systems might ideally present important novel characteristics (separate from the performer) that might become apparent during the musical performance. Here, Paine considers the human *conversational model* as an ideal. The brilliance of his approach is that it specifies a social communication model as best suited within the context of developing spontaneous inter-agent relationships during a live musical event or interactive experience. He recognises that the presence of living human beings (and perhaps computerised equivalents) bring with them an extended collection of acceptable (and not so acceptable) *behaviours* that directly concern both the audience and the performer(s). To add weight to Paine's proposal, we only need to consider the wealth of social interactions that occur in chamber music and/or improvisatory performance contexts. Here, even small variances the behaviour of performers can impact on the quality and success of musical outcomes.

Although this topic is far too complex to address in this book, it is important to acknowledge it as a potential model, and an alternative to the more dominant *digital instrument* designs.

Gestural Artmaking Strategies

In the process of applying Emmerson's elaborated model with reference to Wanderley's and Paine's supporting models, several fine-grained details come to light. Examples are used to illustrate *event* and *landscape* relationships.

The Event

The event presents a spatial construct that coincides with the human orbit of the solo performer. It encapsulates both the performer and their instrument and would normally not extend spatially beyond the extension of their limbs. It is the birthplace of on-stage musical gesture and live sound and is the place from which the music spreads out – sometimes combining with other performers or simply spreading into the field. As noted, the ideal *event* may be exemplified by any acoustic instrument that presents gesture and sounding result simultaneously. For example, a snare drum hit. To expand the relationship, I offer the example of an amplified snare drum with digital reverberation added. In this scenario an electroacoustic counterpart (of sorts) is created that could not exist acoustically in the performance space but nevertheless remains plausible as it is based on general laws set by the physical universe. It also provides a very clear point of departure from which the musical might begin to spread out further. If we now imagine a pitched sound emerging from the reverberation (in combination with the acoustic snare drum sound) we are presented with a slightly more impossible scenario (the sound of a snare drum, reverberation, and synthesised pitch presented simultaneously), but it is one that remains relatively plausible due to the governance of natural laws observed. A good performer may even adapt the perceived energy of their gesture(s) to more precisely match the sounding result.

This *shadow* or *double* of the acoustic performance provide a basic example of the emergence of what this chapter terms a *digital doppelganger*. The term *Frankenstein(s)* has also been used by various members of the international sonic arts community to describe such plausible yet impossible sonic creations. The term *doppelganger* presents the concept of the *mirror image* - something that appears to be same but may also display subtle differences.

Light Music (2004) – Thierry De Mey

A convincing example of this type of sophisticated *digital doppelganger* functioning within a sonic artwork may be found from 14.20 to 15.40 in Thierry De Mey's *Light Music* (2004) Here, low latency sensors (accelerometers) on De Mey's hands provide access to a range of percussive samples, creating the impression of precision hand movements that are able to *create* a series of expressive rhythmic sounds. To solidify the illusion, De Mey makes use of one very special sound: *the clap* (caused by clapping his hands together). The presentation of this overt isomorphic relationship and naturally occurring schema is essential in ensuring the plausibility of all other fabricated movement-to-sound relationships employed. The *clap* presents a mirror image of his acoustic performance, enabling and supporting the extended sound palette that is used in combination with other expressive gestures.

Light Music presents an interesting observation – that some gestures offer stronger sounding cues than others. The clap, for example, offers a natural (acoustically supported) gesture-to-sounding-result relationship that can plausibly be extended to a wide range of sounds, while other gesture/sound combinations require more imagination to reconcile. This gesture-to-sounding-result relationship, (previously been described as *isomorphic*) not only *forms* the impression of local activity, but also provides a point from which event relationships may be extended (through stage, arena, and landscape spaces) at first becoming stretched

gestures (maintaining at least onsets) and eventually transforming into *concomitant* relationships conscripting other imaginary agents from the field. The concept of *digital doppelganger* is a useful one, in that it clearly describes this point of departure.

Me the Machine (2012) – Imogen Heap

An example that extends the gesture-to-sound relationship in this way, is drawn from Imogen Heap's *Me the Machine*, presented at *Wired 2012* in which she adopts a number of interesting performative strategies that act to help form such relationships. At 13.20 she begins the work by waving her arms in the air above her head. This relatively large gesture to small sound relationship suggests the disturbance of air (or perhaps water) and provides the rhythmic underpinning for the work. At 13.55 she slaps her thighs repeatedly to access tuned percussive sounds, then later at 16.10 extends these gestures to *air drumming,* accessing two tenor drum sounds (marching drums). At 14.55 the combined velocity of her right hand movements are used to overtly manipulate the pitch and timbre of sound-based textures - then from 15:10 to 16.10 the height and position her left hand is used to access a range of discrete pitches (synth sounds). In the given schema, height corresponds directly to high-low sounding results, but interestingly, she accesses the lowest note in the series by fully extending her arm to the left (a pragmatic concern perhaps). Throughout the performance she also uses both her hands as well as her on-stage body position to trigger and control a variety of sounds, loops, and advanced effects including vocal harmonies and reverberation. The success of the work, and its point of difference from other works in its class can arguably be attributed to the overt cause-and-effect (isomorphic) relationships it demonstrates, as well as the way in which it *extends* these plausible relationships to incorporate a range of novel sounding results.

Me the Machine presents several specific gesture-to-sounding-result techniques that are worthy of discussion within the current context. From 14.55 to 15.25 both sound-based and lattice-based mapping strategies as discussed by Paine (2003) are employed. The differences are distinct and striking – supporting many of the assertions he makes. Whether or not the gesture-to-lattice relationship might be perceived as a 'category mistake' remains somewhat speculative, but the ability to compare both strategies within the same work certainly offers food for thought. The *gestural agents* employed in the work include the performer's voice, hand, and body movements, but also a range of cyborg-like machine agents (as is suggested in the title of the work) whose energy is derived from sources *other than* the performer's human gestures. An example of this type of interaction *between* performative agents (Imogen and her machine) occurs briefly in isolation at 14.55 and again at 16.47. Here, and at several other moments throughout the work *glitch* materials (depicting computer/machine-like agents) are manipulated by limbs and hands, which form relationships with her body but also remain somewhat separate from the acoustic performance. Other moments demonstrating similar dualities include the control of pitched synth materials from 15.10 to 16.10. Here, bodily movements are perceived as responsible for pitch changes but not for maintaining the continuity of the notes. For most listeners, this constant energy is more likely to be attributed to an alternative power source (such as *electrical*) whose origin lies *elsewhere.*

Of special mention is the use of *loops* in the work. On at least two occasions (14.03 and 16.27) cause-and-effect gesture-to-sound relationships are deliberately suspended and sounds are immediately transferred to the domain of accompanying mediated materials. This sample-and-loop strategy is extremely common and forms the basis of all works that employ *loop station* technology. For many listeners/viewers this approach is simply part of the *allowable* range of techniques supplied by the technology, but for others the effect can be somewhat jarring, drawing unwanted attention towards the technology itself and in the process unhinging the plausibility and communicability of the work. Our discussion to date suggests that any such change in relationship (in this

case from isomorphic to concomitant) might ideally be handled with care and rigour, and that justification of technique might be considered in these circumstances. The *loop station* concept also tends to bring with it certain expectations in terms of structure. In such contexts, a listener/viewer might predict the likelihood of *layering* (of tracks) leading to the emergence of certain types of structural forms.

The Stage

The stage is an area where interaction with other performers occurs. It is ordinarily populated with real performers (other musicians) who are present in the space, but it can also contain imaginary performative agents who are *perceived* as other performers. Such is the case in live sonic arts scenarios where interactivity between performer(s) and the computer is the focus of creative exploration. In 2020 an extended practical study was undertaken at the University of Auckland to determine the point at which a live solo performance might be perceived as an interactive performance (involving two or more gestural agents) with the aid of computer technology. The study employed a traditional side-blown flute, a microphone, and software that precisely measured the attack, sustain, and release (ASR) of the volume of the instrument, as well as the pitch (measured in hertz). This information was converted into MIDI data and transferred (in real time) to physical modelling software that provided a convincing double (*doppelganger*) of the acoustic flute sound. As part of the study, a simple solo melody was played, and the *doppelganger* was manipulated to gradually depart from unison, eventually providing harmonic accompaniment and alternative melodic lines. The following results were obtained through a response survey in which several postgraduate music students participated:

- Maintaining the input ASR curve and pitch data resulted in the perception of one audio stream
- Time variations in the release of the amplitude envelope formed the first (subtle) level of departure
- Time variations in the sustain of the amplitude envelope formed the next (more obvious) level of departure

- Time variations in the attack of the amplitude envelope formed strong levels of departure and impressions of decoupling from the accompanying acoustic sounds
- Variations in pitch (for example harmony in thirds) formed strong impressions of departure but not as strong as variations in ASR attack.
- Time variations in the attack of the amplitude envelope together with variations in pitch (for example harmony in thirds) formed the strongest level of departure and impressions of decoupling from the accompanying acoustic sound

The impression of the emergence of dual gestural agents at the junctures indicated was aesthetically striking and artistically potent. The study proved to be useful as it not only provided a ranking in terms of effective techniques relevant to the practice of instrumental performance with live electronics but also a framework that might be suitable for the future development of artificially intelligent (AI) performative gestural agents.

A theoretical extension was subsequently developed through discussion with participants around the paradigm of musical AI, which was mooted as the ability of a gestural agent to *listen and respond appropriately* in musical ways. It was proposed that the next level of departure might involve *call and answer* algorithms (antecedent and consequent phrases) designed using heuristic if-then statements to access a range of appropriate responses to measured inputs. For example, if a *call* phrase is measured as ascending, then an appropriate *answer* phrase might be to descend in the opposite order of notes. The possibility of measuring (and responding appropriately to) tonic or dominant acting notes and melodic phrases was also discussed. From these humble beginnings, it was thought, we might begin to see the emergence of something resembling the *conversational model* put forward by Paine (2003) as an ideal.

A fascinating outcome of the study was the early point at which independent gestural agents tended to be perceived. Even a basic reassignment of ASR values measured from the flute performance and transferred to the digital instrument in real-time was enough to present

the impression of an intelligent musical accompanist. The observation appears to coincide with Bregman's description of 'acoustic regularities' presented in *Chapter Two: Multimodal Listening*. For example, regularity 1 states: 'unrelated sounds seldom start or stop at exactly the same time' and regularity 4 states: 'many changes that take place in an acoustic event will affect all the components of the resulting sound in the same way at the same time'. It seems that these 'primitive schemas' may be responsible for the *impressions of decoupling* experienced by project participants. Around the same time (2020), a team of Australian researchers: Jon McCormack, Patrick Hutchings, Toby Gifford, Matthew Yee-King, Maria Teresa Llano, and Mark D'inverno established a similar conceptual framework for AI interactivity. They write:

> An alternative is to develop new models of interaction and co-creation that are designed to nurture and enhance the user's creativity and creative practice. In this mode of AI-as-collaborator, human–machine collaborations have shown to foster human creativity in specific contexts (Liapis, Yannakakis and Lopes 2016; McCormack, Gifford, Hutchings, Rodriguez, Yee-King, and d'Inverno 2019). One approach is to parameterise systems for the generation and curation of output. In music composition collaborators specify parameters such as the types of melodies and themes for generated compositions to be based on, time and key signatures to constrain the number of melodies, and so on (Morgan, Ackerman and Cassion 2018; Cassion, Ackerman, Loker and Palkki 2017). Higher levels of autonomy have also been seen in systems that make decisions based on metadata-level criteria. For example, the goal-awareness approach of Hantula and Linkol (2018) models its collaborators by measuring their musical potential in order to make decisions about how to carry on with the collaboration (McCormack, Hutchings, Gifford, Yee-King, Llano, and D'inverno 2020: 42)

Simon Emmerson also offers some useful thoughts on the ideal relationship between musicians and machines in his 2015 *Organised Sound* article *Listening With Machines: A shared approach.* He distinguishes two main types of relationship as follows:

> In discussing how machines and humans interact I have suggested two distinct paradigms of human–machine relationship (Emmerson 2009: 168–71). The first extends our (musical) reach, the expressive possibilities of our

> physical and mental 'event space'. In this view the 'deep' history of our technology in the world has been about prosthesis, both physical and mental – we want to stretch out and contact others way beyond the boundaries of our immediate senses. From memory to writing, then to communications technologies, machines have allowed us to of load and store first plain data, then sound and video, and now processing (and perhaps eventually thinking itself). The second has a different emphasis– the machine as a clone 'other', eventually acting and interacting with us indistinguishably from any other human agent. (Emmerson 2015: 71)

Emmerson then sustains his argument that '*machines are not very good listeners'* proposing an ideal relationship that does not seek to '*off load'* sophisticated tasks such as aesthetic decision-making to machines, but rather to seek out a partnership that is more fruitful. He writes:

> In conclusion, I suggest an approach that encourages the active collaboration and participation of the user with the machine partner, one that enhances and does not off load the fundamental skills of listening. Care must be taken to keep transparent and open to question the complex and dynamic relationship of measured as present to perceived as present. It is clear that robotic assistants (real or virtual) will play a key role in the future of music but their job descriptions, which will no doubt change over time, must be clear. (Emmerson 2015: 74)

The Arena

The arena is the area that exists halfway between the stage and the periphery of the concert hall. It encompasses the audience's seating area and begins to extend outwards towards the surrounding landscape. It is the spatial construct that houses the entire practice of *performing arts* and is the place from within which gestural agents might begin to transition to the landscape - accessing energy sources that are beyond the *performance arena.* In the process of *spreading out,* most gestural agents undergo transitions from *gestural* to *textural* domains. Musical materials tend to begin to become extended in duration and in the process their perceived causes metamorphose from anthropomorphic to *environmental* energies. There are at least two points worth discussing further with reference to this process. The first concerns what Denis Smalley calls *surrogacy.* As sound-based information is reduced (through

stretching or spectral transformation) a sound becomes less recognisable, perceptibly widening the range of its possible causes; however, if that sound is *gestural* in origin it will tend to maintain its association as anthropomorphic even when only highly abstract sound remains. Smalley creates the continuum of *surrogacy* to describe this process that has already been discussed in chapter five. The second point relates to perceived timeframes. As gestures are stretched, they reach a natural threshold where ASR envelopes (amplitude, pitch, timbre, etc) become difficult to follow and where attention begins to be drawn to internal sound qualities. This point of impasse arrives surprisingly quickly. An exercise that demonstrates this phenomenon involves the creation of vocal sound and accompanying hand movement within the gestural timeframe of 3-5 seconds. If the same sound and movement is *stretched* to a longer timeframe (say 12 seconds), it naturally begins to enter the *textural* domain.

The Landscape

The landscape is an area that lies outside the performance arena. It begins on the periphery of the presentation space and expands to the environment, offering interaction of all sorts with the outside world. Performative agents that occupy this space include mediated agents that are *brought in* for various purposes (such as environmental recordings), as well as those who may have started life in the event area yet developed independent agencies their own, eventually spreading out to the landscape space. A common method of expansion from the event (through the stage and arena) to the landscape is the technique I refer to as *field sustain.* Here, performance gestures/sounds enter the arena/environment via a stretching-out of ASR curves. Sounds are still subject to diminuendo but at a much slower rate. Within these scenarios the spatial position of multiple loudspeakers also tends to be important. They must be placed on the periphery, and ideally, a local speaker should also be placed (hidden) behind the performer to present a supporting localised event space image.

The Road Home (2011) - Xinh-Xo Nguyen

An example of the effectiveness of the *field sustain* technique is provided by Xinh-Xo Nguyen in *The Road Home* performed at the University of Auckland as part of the Australasian Computer Music Conference 2011. The work is presented in arch form, with the performer (Xinh-Xo) taking some time to *build up* a rich sound morass on the periphery. The composer creates a scenario where constant energy (performance gesture) is applied - effectively stretching his gestures into continuous textures. This allows materials to *sustain* (dislocated and swirling) for a short time in landscape space. The construct presents the idea that the environment acts as a kind of *battery* for stored and dislocated performance energies. The resulting gestural agents are perceived as fragile and expressive as they are only able to exist for relatively short timeframes before losing energy. An interesting aspect of Xinh-Xo's work is that it makes use of a well-known physical model (friction) drawn from nature. It suggests that the landscape is a domain in which these natural laws of physics might apply.

The concept of the *physical model* (defined by *Newtonian laws*) as a means of supporting local-to-field relationships is an intriguing one. The paradigm has been discussed in chapter six with reference to the plausible/natural cause-and-effect relationships is exhibits in the form of audiovisual relationships. Here, it is extended to the domain of *gesture*, which also connotes an element of proprioception or *felt* bodily experience.

Pillars of Introspection (2019) – Clovis McEvoy

A work that extends on this paradigm is Clovis McEvoy's *Pillars of Introspection* (2019) – an interactive work for virtual reality (VR) in which laws relating to cause-and-affect, gravity, collisions, and friction play a major role in forming and developing both audio and visual materials. The work gives rise to several powerful gestural agents that derive their energy from field space (the environment) but whose genesis is clearly in response to expressive gestures supplied by the performer. At 6.40 visual and aural materials representing the *breath of the performer* are presented. These *doppelgangers* quickly fade away and fall to the ground in VR space but at a rate much slower than objects would in the natural world. At 7.00 the performer's speech patterns create three-dimensional *paint splotches* on the periphery of VR space. These objects also begin to fade away but at a slower rate again, overlaying, and filling the surround-space. At 8.15 sung performance gestures give *birth* to pitch-matched performative visual agents that *float* in VR space demonstrating near-infinite ASR envelopes. The effect is one that requires a reassignment of agential energy – transference from the event to the field, where continuous (and perhaps cosmic) energy is supplied by the VR space. By 10.00 an entire *world* has been created in which highly expressive materials musically and visually interact – a world created *through* gestural interaction with VR space. The plausibility and success of the work can be somewhat attributed to the advanced physical models employed throughout, as well as the gesture-to-sound/image pathways that have been managed with care and rigour.

Beyond virtual reality, a few specialised audiovisual modular programming environments provide robust platforms for investigating the merits of physical models. *Jitter Physics* is one that allows the creation of physical *systems* within which numerous objects may interact to supply and adapt both audio and visual materials. Data streams include x, y, z spatial position, gravity, friction, force direction and strength of impacts,

reciprocity (bounce), etc. Although the models are built on visual force systems, they follow basic rules that have been supplied by nature and therefore provide *plausible* development strategies for both sounds and images. The work of Federico Foderaro (federicofoderaro.com) offers a valuable resource to extend on this paradigm. He supplies numerous examples of physical systems including wind, fluid and other natural *simulations* that can be used as *systems* from within which a range of gestural agents (both local and field) might be seen to coexist and interact with one another.

Concluding Comments

The outcome of the chapter has narrow and wide implications. From the narrow perspective of instrumental performance with live electronics, the adoption of fine-grained synthesis techniques (MIDI, physical models, etc) opens the way to achieving express control of gestural-to-sonic relationships that can serve to fabricate accompanying gestural agents in event and stage spaces. These techniques provide a point of departure for moving from the local to the field. Audio processing techniques (such as those examined in *The Road Home*) can provide robust outcomes, particularly within the context of creating gestural agents in arena and landscape spaces. A combination of synthesis and audio processing techniques might ideally provide the best overall results within the context of digital instrument design. In field spaces, universal justification of sonic transformational techniques is available in the form of physical models (drawn from the environment). Many functional relationships between local and field gestural agents appear to be juxtaposed (located in event and landscape spaces rather than in stage and arena spaces). The in-between spaces (stage/arena) appear to provide fertile ground for novel creative exploration.

From a wider perspective, the chapter has emphasised the presence of gestural agents within all genres and categories of music and sonic art. The general concept of *gesture* has been defined and discussed through its expressive (body-centred feelings), visual and spatial dimensions and through its relationship to the environment, but a deeper investigation

reveals that it is entangled with all eight modes of listening identified. The *formal mode of listening* (for example) can be viewed through the lens of image schemas, which have a direct relationship to the human body. It has been suggested that even the most remote of *referential* sounds may be open to anthropomorphisation and appreciation in terms of gestural agency (real or imagined), and to complete the construct, the concept of *structure* may be appreciated in gestural timeframes and larger timeframes including *re-scaled* gestures (as the previous example *The Breaking of The Scream* by José Halac demonstrates). To summarise this most general point, *gesture* may be seen to form a natural component of all modes of listening identified.

Chapter Eight: Environmental Connections

In the context of artmaking with sound, the topic of *environmental connections,* has both wide and narrow connotations. From a wide perspective it embraces the human-nature relationship as a mode of listening. As previously discussed, this way of appreciating sonic materials is naturally entangled with at least referential, visual, environmental, and spatial modes of listening, and can also be seen to gently stir the remaining four modes identified: expressive, formal, gestural, and structural. Of special mention is the relationship it forms with the *gestural* mode of listening in offering a counterpoint to it (in the way specified by Emmerson's (2007) *local-field relationship*). This connection may be further understood through the gesture-texture continuum (centred on expanding timeframes), and the concept of *event space,* which has been described as *nested* within *landscape space*. The environmental mode of listening presents a foreground/background spatial construct that is fundamental to the listening experience - with *landscape space* providing a context or background in which all sounds are located (discussed in detail in *Chapter Nine: Spatial Forms*).

From the widest perspective, environmental spaces are fundamental to the survival of our species. Places of safety and belonging evoke the concept of *home* (going away from, coming back to, etc.)*,* which is deeply embedded in the human psyche. There is also some evidence to support the theory that the auditory process itself developed in response to certain regularities provided by the natural environment – an evolutionary process that has been referred to as *psychophysical complementarity.* The proposal conceptually overlaps the *Savannah Hypothesis* (concerning the evolutionary development of bipedalism) but differs from it in that certain *regularities* can be proven to exist (see Shepard 1981, Bregman 1993, et al). This significant relationship has already been highlighted as archetypally active within all eight of the modes of listening identified.

From a narrow perspective, there are several specific audio-based creative practices that interface directly with the environment – commonly forming part of the practical *work* of composers and sonic artists carried out in the field. These numerous conventions include soundwalks, stereo and multichannel field recording, spectral transformation (the use of software) in the field, loudspeakers in the field, instrumental improvisation in the field, found instruments, environmentally activated instruments (both acoustic and sensor-based), site specific installations, and audiovisual work in the field. This general paradigm, concerning active engagement with, and study of the natural environment is important, as it highlights the fact that sonic artists (particularly) are not confined to the desk or the studio, but tend to be active in the field. This fascinating variety of creative practices will be examined further in the latter half of this chapter.

Background

To initiate a discussion centred on the *environmental mode of listening*, I offer the following comment by Barry Truax. He states:

> Given its by now lengthy history, the idea of listening to environmental soundscapes as if they were music can now be regarded as a tradition (Truax 2008: 106)

This specific mode of auditory attention embraces a range of environmentally led concepts that are not addressed by any other mode of listening. Facets include emotional connections (feelings of awe and reverence), evolutionary relationships (psychophysical complementarity, spaces conducive to human life, etc.), connections to *acoustic ecology* (social, cultural, and scientific aspects of the sonic environment) and the idea of listening *to* the environment (as a metaphor) – that is, measuring human impacts on the natural world and changing behaviours to support reciprocal relationships between the listener *and* the environment. Auditory examples corresponding to this mode of listening include thunderclaps, storms at sea, and wind through the trees. Popular listening practices situated within the natural environment include *deep*

listening and *soundwalks.* Deep listening is described as 'a form of *meditation* where attention is directed to the interplay of sounds and silences' (Oliveros 2005: 2). The approach and accompanying series of environmentally based exercises was championed by Pauline Oliveros in her book *Deep Listening: A Composer's Sound Practice.* It essentially concerns the adoption of an aesthetic attitude towards sound and silence that shares many similarities with other soundscape/ecological approaches to sonic appreciation including those proposed by scholarly sonic artists such as Murray Schafer, Barry Truax, and Hildegard Westerkamp. Schafer's book *The Soundscape: our sonic environment and the tuning of the world* (1994) notably champions the popular terms *soundscape,* and *soundmark* (the sonic equivalents of their visual counterparts of *landscape* and *landmark*). He writes:

> Some of the sounds of the farm have changed little over the centuries, particularly those suggesting the commotion of heavy work; and the voices of animals too have given a consistency of tone to the farm soundscape. But there are also vernaculars. From my own youth I recall a few. The first that comes to mind is the churning of butter. As the churn was pumped for half an hour or more, an almost imperceptible change in tone and texture occurred as the slopping cream gradually turned to butter. The hand operated pump, also on the decline, now snaps into memory as a soundmark of my youth, although at the time I listened to it carelessly. There were others too, like the ubiquitous cackling of geese, or the swoosh and bang of the screen door. In the winter there was the heavy stamping of snow boots in the front hall, or the scream of sleigh runners over hard packed country roads. In the silence of the winter night there might be a sudden crack as a nail sprang from a board in the intense cold. And there were the deep pedal tones that came again and again in the chimney flue during night winds. Then there were the regular rhythms like the gong which brought us in for dinner, or the whirring of the windmill, which the women put in motion at four o'clock each day to pump water for the returning cattle. (Schafer 1994: 49)

The extract presents an alternative listening strategy to those proposed by Schaeffer (1966) through highlighting the human-nature relationship as the basis of what was later termed *soundscape composition*. Truax (1984) goes further still in articulating a collection of three listening

modes based on the human ability to attend and probe environmental spaces (also described in chapter one). He proposes:

- Background listening – the awareness of sound is present to our consciousness (and may persist for a while in memory) but has no specific impact or significance.
- Listening-in-readiness– specific significant sounds, if and when they occur, will provoke conscious attention (even if we are asleep).
- Listening-in-search– we consciously and continuously listen out for specific sounds, characteristics and qualities. (paraphrased by Emmerson 2015: 71)

Truax augments this continuum (in both directions) through the employment of audio technology, adding the following modes of listening:

- Distracted listening (e.g. habituation to media and music as environment)
- Analytical listening (e.g. the discernment of sound qualities, good and bad reproduction, separable parameters of sound, and so on to spectromorphology) (Truax 2008: 107).

Simon Emmerson points out that in situations where source/causes are 'bracketed-out', the *listening-in-search* mode might naturally extend to encompass Schaeffer's third mode also, involving perusal and enjoyment of sound qualities. He states: 'This is the basis of *soundwalks* (Drever 2009) and Denis Smalley's detailed and meditative reflections on space (Smalley 2007)' (Emmerson 2015: 71). Here, attention is drawn to objects and events that might occur *within* environmental space, and of particular importance is the connotation that certain qualities from the more general human-nature relationship may be retained. John Cage was also an advocate of this way of appreciating sound in general – as it naturally occurs, without the ego or will of the composer placed upon it. His approach is summed up in his following famous quote originally spoken in 1961:

> Nothing is accomplished by writing a piece of music.
> Nothing is accomplished by hearing a piece of music.
> Nothing is accomplished by playing a piece of music.
> Our ears are now in excellent condition. (Cage 1973: 3)

Although Cage's statement may be attributed to the social and political activism of the time, an element might also be assigned to the concept of natural balance and order – reflecting his Buddhist ideologies.

In the following comment, Hildegard Westerkamp touches on similar socio-political issues, notably highlighting a tangible overlap between *soundwalking* and larger life concerns in presenting the indistinguishable literal/metaphorical concept of 'taking the time to listen'. She writes:

> Taking the time to listen goes against today's 24/7 status quo of a hectic pace and stress, of racing toward riches and success, of never having time and always being importantly busy. In this larger context, listening is a conscious practice in learning to change our pace in a society dangerously speeding out of control. Out of that doing comes an entirely new experiential knowledge. This is precisely what makes out the disruptive nature of listening. The decision to do it, rather than talk about it, is the first disruption of old patterns and gives us a visceral understanding of what in fact the disruptive nature of listening really is. (Westerkamp 2019: 47)

Westerkamp also presents strong arguments concerning the necessity of listening to *silences* (both literally and metaphorically) pointing out the *silence* of certain indigenous voices in various socio-political contexts. She asks:

> Is our listening open enough to enter into a true relationship with indigenous people, their cultural and environmental knowledge? (Westerkamp 2019: 61)

Soundscape Composition

Barry Truax combines a number of these important principes based on the human-nature relationship in the following description of *soundscape composition* - a practice that primarily emerged from the World Soundscape Project (WSP) at Simon Fraser University in Vancouver.

> The characteristic principles of soundscape composition as derived from its evolved practice are: (a) listener recognisability of the source material is maintained, even if it subsequently undergoes transformation; (b) the listener's knowledge of the environmental and psychological context of the soundscape material is invoked and encouraged to complete the network of meanings ascribed to the music; (c) the composer's knowledge of the environmental and psychological context of the soundscape material is allowed to influence the shape of the composition at every level, and ultimately the composition is inseperable from some or all of those aspects in reality; and ideally, (d) the work enhances our understanding of the world, and its influence carries over into everyday perceptual habits. Elsewhere I have described the ideal balance that should be achieved in such work as matching the inner complexity of the sonic organisation with the outer complexity of relationships in the real world, without one being subordinate to the other (Truax 1994a). (Truax 2008: 106).

A method of navigating between the ideologies of *soundscape composition* and more practical approaches is also put forward by Truax. His comments stipulate a precise process by which composers and sonic artists may identify and extract objects and events from within environmental space for the purpose of using and developing them in sonic artworks. The process involves a shift in modal listening focus from the largely *reflective* environmental mode to a variety of other modes well suited to *probing* for specific sound qualities (housed by Truax's concept of *analytical listening*). The passage provides a *go-between* - bridging the wide and narrow contexts of this chapter.

> My own practical suggestion with regard to soundscape recording and composition is to begin not with recording or processing in the studio, but rather with the experience of soundwalking in the soundscape. Soundwalking is best done with the only intent being listening, without the distraction of operating a recorder. It is arguably the most direct aural involvement possible with a soundscape and one where repetition does not dull its effectiveness, since each walk is unique and unrepeatable. It is also a good practice to open one's ears and self to whatever is inherent in an environment, with minimal preconception, ideally treated as a phenomenological experience. Several soundwalks in the same area are desirable at different times of day or even seasons, since any one exposure is a small slice of the totality. This advice is as true for one's own community

> as for where one is a visitor (but preferably guided by a local). After various soundwalks have been completed, the listener can make a more informed choice about recording, along with what is the best way to represent the soundscape through recording. Later in the studio, the question as to what a reasonable representation of the soundscape is can be judged based on that experience. This is not to say that an intentional emphasis or exaggeration cannot be contemplated for a specific effect or to raise an issue. (Truax 2012: 196)

The latter part of Truax's comment references a process concerning the operation of audio recorders in the field – a common practice for sonic artists. In a survey of seventeen New Zealand sonic artists carried out in 2010, it was found that fifteen of the seventeen interviewed regularly undertook sound recording excursions in the natural environment. One individual also stated that association with the natural environment was 'what attracted [him] to the domain in the first place'. In the same survey it was observed that other types of *fieldwork* were also routinely carried out within these natural settings. A table displaying the variety of practices associated with the natural environment appears below (table 8.1):

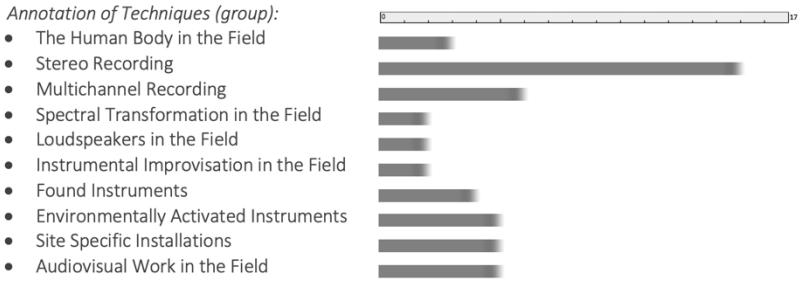


Table 8.1: Collective annotations concerning techniques employed in the field

Notably, excursions undertaken by the seventeen surveyed were not limited to engagement in one or two practices in the field but often involved a variety of artmaking approaches and experiments. To present this aspect of the survey, a representation of the *volume* of work

undertaken in the natural environment is shown in the pie-chart below (table 8.2):

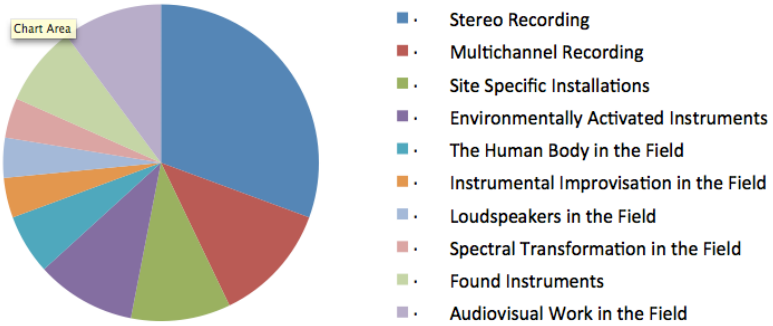


Table 8.2: Volume of field-related activities carried out by the 17 participants surveyed.

Although some of the approaches to working in the field were shared through the availability of supporting technology (for example, stereo, multichannel, and audiovisual recording), It was evident that this constituted only 50% (approx.) of the total work carried out in the field. Several of the seventeen sonic artists surveyed were actively engaged in what they perceived to be their own *unique* methods and techniques of *working* in the natural environment (for example, through *instrumental improvisation*, and the use of *loudspeakers in the field*). Although a small percentage of this work was repurposed for the studio, in a surprising discovery, it was found that almost *half* of the work carried out in the field (by this group) was focussed on working in the environment as a specialist art practice – that is, the majority considered the natural environment as their art presentation space. Furthermore, most of this work was carried out alone – connoting a 'meditative' approach (rather than a performative one) in the manner suggested by Emmerson (2015). We now turn to some of the specific categories identified in the survey as a means of presenting a dynamic range of methods and techniques that may be relevant to composers and sonic artist working in the field.

Soundwalks and Participatory Soundwalks

Although the group surveyed in 2010 did not specifically mention *soundwalks,* it would not be overreaching to suggest that their aesthetic approaches to *work in the field* was very similar to those described by Schafer (1994), Truax (1994a, 2012), and Westerkamp (1974). In the following extract Hildegarde Westerkamp provides comment relevant to several of the practices mentioned (inc. the human body, instrumental improvisation, and found instruments in the field) – effectively combining them under one umbrella of 'participatory soundwalks'. She writes:

> In this context it is appropriate to mention another type of soundwalk which does not only include attentive listening but also active physical participation in the music of our environment. There are many opportunities for this kind of activity in the soundscape, which children often demonstrate to us quite naturally. But to many of us adults, the idea of creating our own sounds, of composing or orchestrating our environmental music may seem silly and contrived. But surprisingly, going on a participatory soundwalk can create unexpectedly interesting dialogues between our surroundings and ourselves. It is worth a try. (Westerkamp1974:3)

Stereo Recording in the Field

Naturally occurring site-specific sounds, such as wind, rain, thunder, streams and rivers, surf, geothermal activity, birds and insects offer huge potential to sonic artists as these sounds present their own referential meanings as well as ideal points of departure towards the musically abstract. Their pitched, rhythmical, spectral and gestural qualities provide a point of intersection with a range of instrumentally produced and/or synthesised sounds that may include found and fabricated instruments made from environmentally sourced natural materials such as wood, clay, stone, bone and shells that borrow their design and noise-producing capacities from nature itself. Sounds of the human voice too, both spoken and sung, might be employed in the field to imbue the sonic textures with the qualities of utterance and behaviour, as well as semantic materials (spoken narratives) as required.

Combinations of these sound categories might also be used to bring to life *inaudible* natural events, such as the setting of the sun, and the feeling of wind against the skin. Digital technology (sensors, real time transformation, etc.) may also be used to augment and extend the intrinsic qualities of what is visually, viscerally or acoustically present in each given natural setting. This imaginative approach to stereo field recording (intersecting with several other associated practices) is presented as an alternative to the *archivist's* approach, which simply aims to re-present the meditative experience of listening in the natural environment. In the view of the author, this is rarely achieved in stereo recording, as there is a disparity between the psychological impression of experiencing sounds, images and feelings in the natural environment, and what can be captured and re-presented using audio recording technology. Focusing on specific objects and events within environmental space seems far more fruitful – an approach that literally evokes Truax's *listening-in-search* strategy.

Carrying out this type of creative practice in the field is not new. Objects and events that constitute the natural environment have been *captured* and *abstracted* by artists (painters, poets, etc.) for millennia. In the latter half of the twentieth-century French philosopher Gaston Bachelard undertook a detailed study with a specific focus on the four elements fire, air, water and earth. In a series of five books, he claims that oneiric (dream-like) qualities are inherent in the material qualities of the elements themselves, drawing on evidence from hundreds of sources (poets, writers, etc.) who imagined identical images relating to these features. His thesis is offered in heavily abridged form here, as it provides a counterpoint to the *listening-in-search* mode of operation under discussion. Bachelard claims that the material qualities of the elements themselves *afford* certain possibilities to the listener/seer in terms of the oneiric *images* they present. The superimposition of this overarching concept changes the *search* parameters from the widest imaginable to a *search for specific qualities* that make themselves apparent through environmental sound materials. The examples Bachelard provides are extremely complex, so may not translate well in this brief description.

Nevertheless, let us use the example of a small stream. If we allow ourselves to dream freely, we might imagine at least *falling water* (laughing, babbling, playing, the song of the river) and *the purity of water* (clear and crystalline). Bachelard argues that the qualities of the water itself present these poetic associations together with a range of creative possibilities that might be drawn from them. Such *oneiric* imagery forms only a tiny part of Bachelard's thesis on water.

Multichannel Recording in the Field

Two main categories of multichannel recording practices have been identified: ambisonic recordings, and zoned field recordings. Ambisonic recording techniques (discussed in detail in *Chapter Nine: Spatial Forms*) make use of specialist muti-capsule microphones (such as A-format first-order tetrahedral microphones) as well as *decoding software* that presents the recorded sounds within chosen loudspeaker configurations as three-dimensional sound images. Without going into technical specifics, ambisonics relies on phase differences between matched, equidistant loudspeakers, and a small central listening location (*sweet spot*). The main benefit of using ambisonic recordings within the context of the current discussion is the sense of *immersion* and enhancement of realism they bring to the simple re-presentation of the meditative environmental listening experience. From a practical standpoint, these three-dimensional recordings may be used as background textures in multichannel loudspeaker configurations, or, where active foreground objects and events are also present (for example, the busy *hutongs* of Beijing), the recordings may supply enough interest to be used in isolation.

Zoned field recording techniques evoke the concept of *distance from me* evident in Emmerson's local/field relationship (2007) and Hall's four psychological zones (1966) (discussed in detail in *Chapter Nine: Spatial Forms*). The method essentially involves the deconstruction of environmental space into various near and far zones, as well as high and low regions. The technique calls for maximum creative imagination. The babble of a nearby stream, for example, may be assigned behavioural

qualities (playful, youthful), and be repositioned in the foreground through close recording techniques. It may be brought closer again with hydrophones, suggesting immersion or *bathing* in the stream, while the rumble of the distant main river might be placed in the far and low zone. Cicadas and birds might occupy the upper spaces, while the riverbed rocks (a non-sounding visual image) might be assigned an appropriate *imaginary* sound in post-production. All zones might ideally be recorded synchronously for later playback as *stems* in a multichannel loudspeaker array or with the intention of creating motion pathways between zones. The recordist might also *act out* some of these ideas in the field (travelling downstream for example). A variation on this highly effective technique is to capture each zone *asynchronously* using only a stereo field recorder with a view to reconstructing the combined three-dimensional sound image later in the studio. Zoned field recording can make use of close, as well as ambient recording techniques and as such can be seen to occupy the middle-ground between stereo and multichannel ambisonic recording methods.

Environmentally Activated Instruments

Two main categories of environmentally activated instruments have been identified: acoustic and electronic. In the category of acoustic instruments, specific designs tend to dominate. Aeolian harps are one such design – played solely by the wind, and dating back to antiquity, these expressive instruments respond to changes in the velocity and pressure of air currents moving across and vibrating multiple nylon strings (previously gut). Each string vibrates in simple Pythagorean motion accessing a range of notes from the harmonic series (8ve, 5^{th}, 8ve, 3^{rd}, 5^{th}, 7^{th}, 8ve, 9^{th}, 10^{th}, etc.). Specific notes are produced in response to wind pressure and the *standing waves* that each string sets up (vibrating in divisions of 1, 2, 3, 4, 5, 6, 7, etc.). As all strings act independently, the overall effect is often *chordal* even if the strings are not tuned. The delicacy, complexity, and variety of sounds produced by aeolian harps together with the fact that they *play themselves* has, over the course of history, given them associations with magic, sorcery, and the supernatural. Other environmentally activated acoustic instrumental

designs include *river harps* (strings played by the flow of water) *singing trees* (tuned pipes/flutes played by the wind), and last but of course not least, numerous and varied types of *wind chimes.*

The second category – electronic instruments, within this context primarily involves the sonification of data collected from environmental sensors. Synthesised sound and music may be produced either in real time, or later in the studio. Electronic sensors may be employed to capture real-time data from naturally occurring events (both sounding and non-sounding) such as wind gusts, summer rainstorms, boiling mud pools, breaking waves and the setting of the sun – which may then be used to synthesise or manipulate sonic materials in real time. A site-specific anemometer measuring wind pressure, for example, might be used to advance through a harmonic series (the higher the wind pressure, the higher the partial). This was the approach taken in the author's environmental installation *Windsong* (2019), which might be described as an *electronic aeolian harp*. Data streams from longer-time-frame natural events may also be used to generate and shape music. For example, the fluctuation of tides, the angle of shadows cast by the sun, changes in rainfall and river flow, the growth patterns of plants, the location and strength of earthquakes, average temperatures over the seasons, and the surface area of polar ice during the summer over the past twenty-five years. Data from these natural events, which normally accumulates over days, months or years, is compressed in much the same way as a time-lapse video - and is presented to listeners in audible gestural timeframes for the purpose of revealing the internal patterns. This allows listeners to more accurately perceive the slow changes that are taking place in the natural environment and promotes a deeper understanding of its cyclical and more extended patterns of change.

Employing electronic technology in this way is not without its problems. Data received from sensors placed in the field can be unreliable, and engineering stable solutions can be enormously time consuming. Furthermore, the reliable data that *is* produced often tends to be too basic to use as the only source for synthesis. To create convincing and

expressive synthesised sounds it is often necessary to 'add in' nuance-level changes that sensors are not capable of measuring. This results in a significant element of fabrication that upsets the 1:1 nature-music relationship that is under exploration. Use of sensor technology seems to be a trade-off between these undesirables and the convincing way in which precise electronic measurements expose/highlight a particular element of nature.

Spectral Transformation in the Field

A variant on the sound-producing capacity of sensors in the field, is the sound-filtering capacity of spectral transformation processes in the field. To briefly summarise the approach, the following example is offered: A friend and colleague recorded the sound of an ocean 'blowhole' while on holiday with his family on Fraser Island, Queensland, Australia. When he returned to the studio, he filtered the sound (in real time) to expose a changing pitch series (based on the harmonic series) that was obviously produced by the pressure of the ocean waves through the blowhole. The resulting *music* was both expressive and revealing of detailed information contained in the natural event. Use of such technology in the field might ideally allow processes of this type to be carried out in real time in combination with the natural event itself. Of course this would involve the use of microphones, computers, software, headphones, etc. in the field. Interestingly, as cumbersome as this general method may appear, it was precisely the approach taken by at least two of the seventeen sonic artists surveyed in 2010.

Summary and Conclusion

The paradigm of appreciating the natural environment through the act of concentrated listening has been championed by numerous scholars including Cage (1973), Emmerson (2015), Oliveros (2005), Schafer (1994), Truax (1984, 1994, 2008), and Westerkamp (1974, 2019). Their work has coincided with the emergence of technology that allows this reflective experience to be augmented for the purpose of exposing hidden characteristics of natural systems and events. This is highly significant as it allows the possibility of artificially transforming the human-nature

relationship from observational to collaborative - allowing the will of the participant (their listening attitudes and the technologies they employ) to play a part in forming the experience itself. A very basic example of this principle at work (using Truax's terms, transitioning between *listening in-search* and *analytical listening*) is the use of a microphone and headphones in the field. The audio recording technology amplifies and exposes certain characteristics, encouraging further probing of sounds in very specific ways (a creatively engaging process in itself). The self-referential approach of using exposed natural characteristics to *further expose* those same characteristics (evident in both recorded and data-driven approaches) ensures that techniques and technologies remain as transparent as possible - in that it is the qualities within *the sounds themselves* that justify employment of developmental techniques. This approach - considered highly relevant within the context of *environmental sound,* is also known as 'abstracted syntax' (Emmerson 1986). The latter part of this chapter encourages augmentation of this *listen-in-search* approach to engaging with the environment through the active imagination of the recordist/sonic artist and extension of the technologies they may employ in the field. This narrow scope of the chapter is based on the range of music-making activities that were carried out in the field by seventeen New Zealand sonic artists surveyed in 2010.

From a wider perspective, the chapter highlights an inborn mode of listening that evokes an aesthetic appreciation of the natural environment, extending the musical paradigm through its reliance on the human-nature relationship. This listening attitude has been termed the *environmental mode of listening*. Its operation is a highly significant proposition, as it relocates core aspects of the musical experience within the confines of a deep archetypal appreciation of nature. The phenomenon also has multimodal dimensions and may be further described through each of the eight modes of listening identified. The concept under discussion is one that concerns an inversion of hierarchy from the *environmental mode of listening* as a constituent, to one that repositions *the environment* as entangled with, and central to, all other

modes. Barry Truax presents this principle (of inversion) in the following comment. He states:

> I would now like to suggest inverting this increasingly familiar concept to suggest that we listen to electroacoustic music as if it were a soundscape. (Truax 2008:106)

As previously mentioned, environmental associations are heavily entangled with at least referential, visual, and spatial modes of listening, and can also be seen to gently stir the remaining four modes: expressive, formal, gestural, and structural (identified in the eight entangled modes model). For the purposes of this discussion, and as a way of further evidencing some of the main points presented, it may therefore be useful to consider the environmental mode of listening in two-way combination with the remaining seven modes. This allows each constituent quality to be specifically addressed. The following synopsis attempts to highlight these most common associations.

The *environmental-expressive* blend of listening modes is interesting in that it asks, *what does nature express?* Such has been the focus of many philosophers and poets throughout the ages. Nature has often been described as 'in a state of grace' emphasising the point that it does not *feel* as human do. However, there is a range of emotions humans commonly experience when *in* nature. For example, the emotions and feelings associated with the act of watching a sunset. At the very least, the *environmental-expressive* mode highlights the human-nature relationship and its emotional complexity.

The *formal* qualities of the environment are extremely complex, reflecting a glittering array of constants and seemingly chaotic patterns produced by its constituent objects and events. 'The butterfly effect' (flapping its wings on one side of the globe and causing a storm on the other) reminds us that these patterns are in fact not random, but rather in a state of perfect balance with one another. This phenomenon was explained to me as a student through the example of a stone on a beach with a sand pattern around it. It was emphasised (by my then teacher) that the kinetic energy of every molecule of water was in perfect balance

with each grain of sand displaced by the presence/position of the rock. The paradigm was held up as a compositional ideal that may be conceptualised as an *environmental-formal* blend.

Environmental sounds are also known for their ability to evoke associative meanings. A river-scape, a train yard, a reverberant space, or the sound of a large crowd protesting stirs range of subjective interpretations in listeners that can activate schemas stored in long-term memory. Such modal listening strategies may be referred to as *environmental-referential* blends. The given schemas (memories), also tend to house *visual* snapshots recalled from life experiences and so may also be considered *environmental-visual* blends.

Interestingly, most environmental energies are defined by their long cycles (change of tides, day-night etc.) but those that are not, tend to capture human attention fairly easily. A breaking wave, or a gust of wind, for example, may be appreciated through its *gestural* content – anthropomorphised to such an extent that its common description is often metaphorical. For example, 'the waves *embraced* the shore', 'the wind *whispered* softly' or '*howled* violently'. Such affordances might be considered *environmental-gestural* in quality.

The remaining blends *environmental-spatial* and *environmental-structural* might be considered in combination. Trevor Wishart (1986) sought to make use of environmental spatial constructs such as foreground-background not only for their unifying qualities but also as methods of structural development applicable in the organisation of both referential and abstract sounds. Barry Truax's (1994, 2008) description of soundscape composition is also highly relevant here, as his methods of *fixed* and *moving spatial perspectives* (covered in *Chapter Ten: Spatial Concepts*) directly present a range of spatial/structural developmental possibilities. For example, a listener's *fixed* point of view (POV) might be explored in combination with *moving* perspectives based on known environmental spatial constructs such as the opening and closing of doors or following the topology of a river from its source to the ocean.

Chapter Nine: Spatial Forms

In 2007, Denis Smalley published a monumental article in *Organised Sound* entitled *Space Form and the Acousmatic Image.* In it, he *re-imagined* the last of his nine indicative fields - *space* – redefining it as a container of the eight other fields: gesture, utterance, behaviour, energy and motion, object/substance, environment and vision. He states:

> But possibly the most significant shift in my thinking is that rather than being the final frontier of investigation, space should now move to centre stage to become the focal point of analysis of, or commentary on, acousmatic music. (Smalley 2007: 54)

His article not only redefined three decades of internationally renowned research concerning *spectromorphology* (a term invented by Smalley) but also assisted in elevating the genre of *multichannel acousmatic music to* a special artmaking category where sound alone (in the absence of any visual data) might be imagined as a *constructor* of such space. At the time of publication, there was great excitement within the international sonic arts community concerning the emergence of this new category of music*,* as it seemed to offer unparalleled potential for creative exploration of the spatial paradigm. Smalley refers directly to this potential as follows: 'Acousmatic music is the only sonic medium that concentrates on space and spatial experience as aesthetically central' (ibid). Peter Bachelor encapsulates the essential role space can play in acousmatic music in his 2015 Organised Sound article 'Acousmatic Approaches to the Construction of Image and Space in Sound Art'. He states:

> The use of the word 'image' implies something experienced as static. Indeed, Denis Smalley suggests that the acousmatic experience of a given landscape builds spatially as we aurally explore its terrain, yielding a complete 'image', which exists outside the bounds of time. Time effectively becomes space (2007: 37–8). For the listener, this image therefore persists as something akin to the memory of a photographic still – an encapsulation of the experience in a single moment. It is the combination of a series of 'stills' that together form the narrative experience of a work. (Bachelor 2015: 150)

We have also seen throughout this book; the important role *space* has had to play in the way in which humans appreciate sound and music in general. Spatial constructs such as foreground/background, high/low, and the general paradigm of motion, provide the frameworks for ways of understanding not only sound, but form the basis of our psychological and physical experience.

These introductory comments provide the narrow and wide contexts for this chapter. The wider concerns relate to the cognitive psychology of human audition – how we perceive and use spatial constructs to make sense of sound in general, and in a wider sense again, how we make sense of the world. Multichannel acousmatic music, by comparison, provides a special narrow context for the creative exploration of such principles in a practice that has come to be known as *spatial composition.* The title of the chapter *Spatial Forms* is one that connotes structure and form relevant to both wide and narrow contexts. It proposes the construction of a theoretical model based on a finite set of archetypal spatial forms that participate in the organisation of sound and in the structuring of artworks (as well as life in general). Smalley (2007) defines the term *spatial form* as follows:

> A smaller or larger spectromorphology or musical passage (but not necessarily an identifiable 'unit') that may be analysed according to its various perceived spatial attributes. Space-form is made up of spatial forms. (Smalley 2007: 56)

Spatial Forms

To get a clearer understanding of the many different types of *spatial forms* and their functionality, we might first imagine the listening position as a central point in spherical space. This results in an appreciation of spatial forms in terms of their perceived three-dimensional location, and more precisely, in terms of their spherical coordinates (distance, elevation, and azimuth). For the purposes of this discussion, these three intersecting values will be considered: *perceived distance from the listener*, *vertical position*, and *surround location and imaging*. Although the three groups do not act independently, they provide useful

categories for contemplating different types of three-dimensional sonic images: near/far, up/down, at specific location on the periphery, and/or spread throughout surround space. The first part of the chapter addresses the profound psychological effects of such spatial forms on human beings and provides a detailed list of the known spatial forms that participate in the organisation of sound and in the structuring of artworks. In the second part of the chapter, the topic of spatial artmaking strategies within the narrow context of multichannel acousmatic music will be examined.

Distance From the Listener

Within this category we find the subtopics of social zones (Hall 1966), local/field relationships (Emmerson 2007), comingness (Lennox, Vaughan and Myatt 2001), and intimacy and immensity (Bachelard 1964, et al). This spatial dimension – also known as the *z-axis* is essentially egocentric, as it spreads out from the listener's head and extends to the horizon. *Distance from me* is the key concept under investigation.

Hall's Four Zones

In 1966, Edward T. Hall identified four spatial constructs central to the way in which humans experience the world. In his book *The Hidden Dimension* he articulates these *zones* as: intimate distance, personal distance, social distance, and public distance. In the following paragraphs I have taken the liberty of freely elaborating his comments from pages 116-127. Quotation marks have been used to highlight direct references.

Intimate distance relates to settings where the presence of other human beings is unmistakable and may be overwhelming due to sensory inputs that have been elevated through close proximity. This information includes sight, smell, body heat, sound, feel of breath, etc. Intimate distance is the space occupied by lovers – inside arm's length, where vocalisations are largely involuntary and are often reduced to *grunts* - where 'a whisper has the effect of expanding the distance'. Hall adds that intimate distance may be forced on individuals in various situations – for example, in confined spaces such as elevators. He states that 'subway

riders have defence devices which take the real intimacy out of intimate space in public conveyances.'

Personal distance is a space normally occupied by friends and family and is measured from arm's length to the 'far phase' where 'two people can touch fingers if they extend both arms'. Within this zone, the three-dimensional quality of objects is particularly pronounced. Hall states that 'objects have roundness, substance and form unlike that perceived at any other distance'. Surface textures are also conspicuous and are easily distinguishable from one another.

Social distance is the space where most casual social, and small-group educational information is delivered and received. Hall measures this zone by stating that 'nobody touches or expects to touch another person'. He also states that 'conversations can be overheard at a distance of up to twenty feet' (6m) – delineating the extreme 'far phase' of the spatial zone. Interestingly, he also notes that 'raising the voice or shouting can have the effect of reducing social distance to personal distance.'

Hall defines the near phase of *public distance* as 'twelve feet' (3m) - the distance at which 'an alert subject can take evasive or defensive action if threatened', and the far phase as the distance at which vocal communication begins to break down. With this noted, some overlap between social and public zones is evident. He also notes that 'linguists have observed that careful choice of words and phrasing of sentences as well as grammatical or syntactic shifts occur at this [public] distance'. Examples of *public zones* range from lecture theatres to large-crowd public address scenarios. At this distance social relationships are no longer subject to the constraints of closer distances where eye contact plays a major role in communication. It is the distance at which the paradigm of *speaker-receiver* emerges - where members of the audience are free to act in a variety of ways that are somewhat detached (although not completely) from the presenter.

Emmerson's Local/Field Relationship

In *Chapter Seven: Gestural Associations,* there was much discussion concerning Emmerson's local/field relationship. It is evident that there are numerous parallels between his and Hall's observations. This is important to note as it places the genesis of all music (bodily in origin) at the centre of the spatial paradigm currently under discussion. To clarify, as previously discussed, Emmerson claims that music *spreads out* from this central point occupied by the performer, while Hall's thesis puts the *listener* at the centre of the spatial paradigm. Interestingly, this type of superimposition is one that tends to occur naturally in the minds of listeners in concertised settings. The phenomenon of projecting *the self* into the centre of the event space becomes evident when audience members begin to *mirror* the actions of performers (facial expressions, mouthed words, bodily tensions and moments of release, etc.). The way in which both performers and listeners' spaces expand from the intimate self to the environmental appears to be largely indistinguishable. *Distance from me* is the uniting concept.

Comingness

As discussed in *Chapter One: Modes of Listening*, from an evolutionary perspective, approaching sounds that cross the boundaries of Hall's four zones are likely to be perceived as the most attention-grabbing of all. To reiterate, Lennox, Vaughan, and Myatt (2001) claim that *rushing sounds* are the most arresting of all sonic materials. Furthermore, they prove through a series of psychological tests that the order in which sonic information is processed is: 1) the quality of *comingness* 2) what it is, and 3) where it is. Their research encapsulates the concept of survival – suggesting that when a sound impinges on our consciousness, the first quality that is attended is whether or not it presents any threat to us.

Intimacy and Immensity

If we imagine the extremities of the continuum presented by Hall (1966) and Emmerson (2007) from near intimate/event to far public/landscape zones, a powerful juxtaposition of spaces emerges that is central to the psychology of experiencing music (as Emmerson suggests). Gaston

Bachelard (1964) goes further still, in claiming that all works of art are a by-product of the human ability to *daydream* of immense spaces. In *The Poetics of Space,* he states:

> In analyzing images of immensity, we should realize within ourselves the pure being of pure imagination. It then becomes clear that works of art are the by-products of this existentialism of the imagining being. In this direction of daydreams of immensity, the real product is consciousness of enlargement. We feel that we have been promoted to the dignity of the admiring being. … As soon as we become motionless, we are elsewhere; we are dreaming in a world that is immense. Indeed, immensity is the movement of motionless man. It is one of the dynamic characteristics of quiet daydreaming. (Bachelard 1964: 184)

Bachelard attempts to reconcile the two extremes of the intimate-immense continuum through the identification of common traits. He states:

> It would seem then, that it is through their 'immensity' that these two kinds of space – the space of intimacy and the world space blend. When human solitude deepens, then the two immensities touch and become identical … The *being-here* is maintained by a being from elsewhere. Space, vast space, is the friend of being. (Bachelard 1964: 203)

Bachelard's poetic descriptions of intimacy and immensity are useful in the current context as they support aspects of discussion concerning multimodal blends between expressive/gestural and environmental/spatial modes of listening - including associations with reflective listening attitudes and commonly experienced feelings of awe and reverence.

Vertical Position

High and low sounds and motion between these extremities (also known as the *x-axis*) presents the next major category for discussion. Sounds that occupy this *above-ear* spatial domain are rare. Beyond birds, the occasional aircraft, and wind through the trees, there is not much that occurs naturally - and this tends to inform what might be convincingly

presented as *situated above*. Notwithstanding these limitations, there are powerful abstract concepts that play roles in the plausible presentation of such materials. These models include *Snyder's Up and Down Image Schema* (2000), *Smalley's Pitch Space* (2007), and *Bachelard's Verticality* (1964)

Snyder's Up and Down Image Schema

Snyder's five image schemas - up and down, centrality, motion- linkage-causation, linearity: pathways and goals, and containment: inside and outside have already been examined at length in *Chapter Four: Formal Musical Qualities.* The current context of discussion concerns the first of the schemas - *up and down*. This spatial construct may be applied to pitch space as well as spectral space, as it is apparent that *high* sounds (of all types) are likely to be perceived as occupying a vertical position higher than accompanying or subsequent *low* sounds. The perceived spatial *range* between high and low sounds may also be presented through materials that naturally contain such dynamic frequencies – that is, the *width* of spatial images (between high and low frequencies) can present the *presentation space* from within which the music/sound might unfold. Beyond musical contexts, the up/down image schema is evident in the ways in which we quantify and qualify all aspects of our daily lives and may be used to describe any increase or decrease in perceived amount or energy or quality.

Smalley's Pitch Space

In his 2007 article *Space Form and the Acousmatic Image* Denis Smalley presents an inventory of attributes relevant to *tonal pitch space.* He references Delalande, who uncovered the following list, which is based on placement in, and motion through pitch space, also taking harmonic factors into consideration. The list includes:

> Transparency of texture, approaching and distancing, 'planes of presence', unveiling, a zoom effect, thickening of melodic line, widening of pitch range; up and down directionality, which includes falls and double falls, upward melodic staircase configurations, poles of attraction, and the feeling of weight. (Smalley 2007: 45)

Notably, *pitch space*, and the vertical axis in general, tend to be described by metaphors that find their basis in observations related to *gravitational forces.* A vivid description of gravitational effects within this context is provided by Huntly (1970). His comments are based on a colloquial understanding of one of the Newtonian laws - all that goes up must eventually come down. He states:

> People had (in ancient times) on unnumbered occasions, observed the flight of a stone in the earth's gravitational field: its rapid ascent, the long pause when it seemed neither to rise or fall, followed by its accelerated decent. Surely the deeply buried memory of this oft-repeated experience is gently stirred when a melody having reached its tonal climax, the high note is sustained for a moment or two before it descends to its level of departure? (Huntly, 1970: 82)

Gravitation

In relation to issues concerning rising or falling, attraction, weight, pitch and spectral space, Smalley quotes Francois Bayle as follows:

> It is Francois Bayle who first called attention to gravitation in the acousmatic image. It belongs to what he calls the 'static archetypes', which define the 'landscape and its population', and include ideas of horizon, temperature, climate. Baylean archetypes, which also include dynamic and position archetypes, are concerned with bases on which existence is founded (Bayle 1993: 76 referenced by Smalley 2007: 45)

Central to Bayle's and Smalley's ideas around *gravitation* is the assumption that the lower regions of both pitch and spectral space provide stability. Through this lens, the notion of *harmonic cadence* (for example) carries with it an associated upper spatial tension that is released by lowering the centre of gravity. Musical terms like *suspension* support this understanding. The notion of *scale* might also be contemplated as a *staircase* with the ground level providing the most stability. As Smalley states, 'the ideas of gravitation and scale' might therefore be considered 'complementary' (Smalley 2007: 47).

Smalley also devotes some time to the notion of 'levitation' noting that 'when levitation in spectral space occurs, we have the impression that

sound floats free of gravitational forces' (ibid). An example of this phenomenon might be a high, strong, continuous tone, usually the tonic - that can act as an upper-level anti-gravitational attractor. Interestingly, materials like this seem to require extra energy to be *thrown* into this upper echelon of frequencies. Smalley notes that it is often a 'coincident or preceding object that launches the projection.' (ibid)

The concept of *spectral weight* is intrinsically bound to the paradigm of *gravity.* It is generally not plausible for sounds containing large amounts of bass to be thrown around at will, or for them to occupy the higher levels of vertical space for any length of time. They tend to move more slowly and normally reside in the lower regions of gravitational space. Sounds that contain higher frequencies are more easily moved upwards – sometimes through their own internal energy, but often in response to other forces acting out on them (for example, a collision). Such details are formed and supported by our innate understanding of universal constants observed in our daily lives. Objects of large-to-small sizes generally correspond with heavy-light weights. Objects of certain shapes and densities also tend to vibrate in certain ways, and these *size/shape* characteristics are reflected in their sound qualities. Objects require energy (either internal or external) to move – especially in the upper direction. If sound objects are perceived as *up*, they will tend to come *down* in a process of losing energy and coming to rest.

Bachelard's Verticality

In *The Poetics of Space* Gaston Bachelard presents verticality in terms of its psychological value. He states:

> Up near the roof all our thoughts are clear. In the attic it is a pleasure to see the bare rafters of the strong framework. Here we participate in the carpenter's solid geometry…. As for the cellar, we shall no doubt find uses for it. It will be rationalized and its conveniences enumerated. But it is fort and foremost the dark entity of the house, the one that partakes in subterranean forces. When we dream there, we are in harmony with the irrationality of the depths… The creatures moving about the cellar are slower, less scampering, more mysterious. In the attic, fears are easily rationalized…' (Bachelard 1964: 18)

This poetic description provides a useful framework for considering the ideal vertical placement of sounds with regard to the referential meanings they may elicit. The passage also supports the *gravitation* paradigm, especially concerning *weight* ('slower, less scampering') and the conservation of momentum.

Surround location and Imaging

This final category relates to the spherical coordinate *azimuth* (angle from central position). It concerns the perceived surround location of the sound object, and well as qualities relating to its three-dimensional sound imaging that may extend from a discrete position to 360º (for example, a 1st-order ambisonic environmental recording). The terms *discrete position, three-dimensional imaging*, and *envelopment* are used to describe the sonic characteristics relating to surround space.

Discrete Position

To gain an understanding of the functionality of sounds that are locatable in surround space, we might first consider their psychological importance in terms of the *spatial* messages they contain. We have previously discussed Lennox, Vaughan and Myatt's (2001) stage 1-2-3 processing model: 1) whether a sound is identified as approaching 2) what it is, and 3) where it is. The current discussion is focussed on category 3 (where it is), and the more general issue relating to the mitigation of *threat.* For example, sounds located *behind* the listener may create a greater sense of unease, while sounds in front of the listener (within the visual periphery) may be *attended* with greater feelings of safety (comparatively). Head turns also form a natural part of listener's responses in 360º listening environments.

Three-Dimensional Imaging

With reference to the term *imaging*, audio is often considered the *poorer cousin* of visual perception; however, researchers have shown that this is not always the case. In the same 2001 article, Lennox, Vaughan and Myatt discuss the specific qualities of three-dimensional information that is most easily 'rendered' in sounds. Their comments (also presented in

Chapter One: Modes of Listening) appear below:

> The 'what' representation yielded in the audio mode is different from, complementary to, and as richly detailed as that derived from the visual mode. Whilst form is not as well rendered, material composition, structure, type-of 'what' and size-of 'what' often are. (Lennox and Myatt 2001: 4)

These comments together with Hall's observations concerning three-dimensional qualities within the *personal zone* provide food for thought concerning the types of sounds that might communicate well through their spatial qualities. According to Lennox and Myatt, *Shape* (form) is not well represented, while *size* (often corresponding to bass content) often is, along with the substance and texture of the object.

Heterotopias

According to Sally Munt (2002), spatial images also present a range of psychological impressions that are central to our daily lives. In *Framing Intelligibility, Identity, and Selfhood: A Reconsideration of Spatio-Temporal Models* (2002) she states:

> Spaces are not only gendered, and sexed, they are also moralised. Crucially, territorial activity is established in the way we live our lives, not just in the political grand gesture but also in the minutiae of existence, and affective relations. (Munt 2002: 2)

Munt goes on to present her proposal that *heterotopias* are the spatial constructs responsible for governing our everyday lives. These 'heterotopias' she claims, 'are kinds of mirrors to utopias, a counterpoint of the real to the unreal, in which the utopic glance returns to reconstruct the real, in a new way of seeing' (Munt 2002: 1). Underpinning her thesis is the tested hypothesis that the process of daydreaming not only forms *elsewhere* spatial scenarios but also creates territories that might be defended in various ways. It is a process that is essentially about 'animating a kind of optimism' (ibid) - allowing latent futures to coexist within the present. The term *heterotopia* has complex meanings and qualities that Munt articulates as follows:

> Heterotopias are the conceptual space in which we live, six principles of which are identified by Edward Soja: i) they are found in all cultures, although no type is universal; ii) they change over time, they have a genealogy and well as a geography; iii) they are dynamic places where many different, incompatible spaces/sites may intersect; iv) they can occur at particular temporal and spatial axes; iv) they presuppose systems of territorial opening and closing which makes them subject to the disciplinary technologies of power; vi) they function in relation to the spaces around them, either through the creation of an illusion or out of the sensibilities of compensation. Heterotopias allow a slippage of meaning, they make it impossible to name this and that... because they shatter or tangle common names, because they destroy syntax in advance, and not only the syntax with which we construct sentences but also that less apparent syntax which causes words and things... to "hold together..." [Heterotopias] desiccate speech, stop words in their tracks, contest the very possibility of language at its sources; they dissolve our myths, [and] sterilise the lyricism of sentences. (Munt 2002: 4)

Complexity is added when Munt proposes that these heterotopias, each of which contain numerous internal components, may also be regarded as multiple versions of *the self.* It is in this way that contradicting values and beliefs collide and *compete* with one another for territory within our own constructed egocentric worlds. Munt refers to this type of interaction as *agency* – noting that it is an essential means of improving the self. She states:

> The opportunity to transform the self is achieved through a network of relations; oneself is positioned in reciprocity to many other selves, some more influential than others, all circulating within certain fields of force. Thus, through interaction and repetition, incitation and struggle is continuously a possibility; in spite of -- because of -- identity ascription, there is always the prospect of refusing what we are. (Munt 2000: 5)

Here, Munt extends on the work of Foucault who first embarked on the project of redefining the human as a spatial invention – a thesis that locates *space* as the prime field – a container of all and everything necessary for the wellbeing of the human psyche. The construct also relates explicitly to the concept of *conceptual blending* previously discussed – as Gary Kendall points out:

> Fundamental to blend theory is the concept of the 'mental space', originally introduced by Fauconnier (1994) and extended by Fauconnier and Turner (2002: 104), who say, 'A mental space consists of elements and relations activated simultaneously as a single integrated unit. Often, a mental space will be organized by what we have called a conceptual frame. (Kendall 2014: 197)

Kendall also offers a description of how these *mental spaces* might be presented and received (communicated) through the medium of sound, and more specifically, within the context of listening to electroacoustic music. He writes:

> The deeply meaningful sense of space that is aroused when listening to electroacoustic music has its roots in a lifetime of embodied spatial experience. Beneath the apparent continuity of everyday space are the axes of the body giving structure and context to the experience of spatial events. For example, objects, actions and agents are most easily grasped in terms of the familiar: objects that are manipulated by one's self. It should come as no surprise then that so much of the spatial movement in electroacoustic music is curved like the movements of the body. Then, too, in the context of the body, motion requires effort. The listener's felt experience of spatial motion in electroacoustic music is typically imbued with the flow dynamics of bodily effort (Kendall 2010b: 232)

Envelopment

In their 2007 publication *A perceptual Approach to the composition of Meaning in Artificial Spatial Audio* Peter Lennox and Tony Myatt, present and describe the term *envelopment* as *where the listener feels immersed by the sound field* (paraphrased). The term encapsulates the concept of an authentic three-dimensional acousmatic image that is as convincing as a real-world experience. *Envelopment* can have powerful psychological effects on listeners – coaxing them into a false sense of reality that carries with it the potential to deliver profound memory-related emotions and socio-cultural frameworks. In the same article, Lennox and Myatt estimate that most of the audible information received by individuals in immersive three-dimensional sound environments arrives via reflections. They argue that it is in this way that humans build up their impressions of three-dimensional space. Their research strongly suggests the presence

of both foreground (noise making) objects and background reflective surfaces is *necessary* in the process of forming such perceptual frameworks. Denis Smalley also supports the view that spaces may be attended and built up in this way. He describes what he perceives as a relationship between spatio-morphology and spectro-morphology as follows:

> I use the term *spatio-morphology* to highlight this special concentration on exploring spatial properties and spatial change, such that they constitute a different, even separate category of sonic experience. In this case spectromorphology becomes the medium through which space can be explored and experienced. Space, heard through spectromorphology, becomes a new type of 'source' bonding. (Smalley 1997: 122)

A Summary of Spatial Forms

As a means of summarising the numerous spatial forms we have discussed to date, I would like to return to Smalley's 2007 article *Space Form and the Acousmatic Image,* which offers the most detailed list available in the form of a glossary. Gary Kendall refers to this important document as 'an extensive catalogue of reference frames' (Kendall 2010b: 232)

Agential space	A space articulated by human (inter)action with objects, surfaces, substances, and built structures, etc. Combines with utterance space to create enacted space.
Approach/recession	Typical relational processes in perspectival space, concerned with shifts and movement between proximate and distal space, and/or between periphery and proximate centre. Associated with ouverture/enclosure.
Arena space	A sub-category of performed space – the whole public space inhabited by performers and listeners, within which gestural space and ensemble space are nested.
Automated mode of delivery	Where automated processes and systems are used to deliver perspectival space in public listening.
Behavioural space	A zone of perspectival space produced by the interaction of sounds which, spectro-morphologically and texturally, indicate collaborative, group identity. See also 'signal space'.
Circumspace	In perspectival space – the extension of prospective and panoramic space so that sound can move around the listener and through or across egocentric space.
Circumspectral space	The spatial distribution or splitting up of the spectral space of what is perceived as a coherent or unified spectromorphology.
Containment/transcendence	How far the acousmatic image appears bounded within the physical confines of the listening space, and how far the image

	appears to transcend such confines.
Diagonal forces	The motion of spectral energy towards or away from a spectral region which acts as a plane. Planes can be expressed or implied. See 'gravitation'.
Distal space	The area of perspectival space farthest from the listener's vantage point in a particular listening context.
Distal interpolation	A temporary rupture in ongoing proximate space thereby permitting access to a distal view.
Egocentric space	The personal space (within arm's reach) surrounding the listener.
Elevated space	A high zone in perspectival space. Not to be confused with 'levitation'.
Enacted mode of delivery Enacted space	The active diffusion of perspectival space in public listening Space produced by human activity – a space within which humans 'act'. See also 'agential space' and 'utterance space'.
Ensemble space	A sub-category of performed space – the collective space within which gestural spaces are nested. Can be thought of as a type of behavioural space.
Fixed mode of delivery	Where the delivery of perspectival space is integrated into the format of the composition and is not further diffused via the enacted mode of delivery.
Gestural space	The intimate or personal, source-bonded zone, produced by the energy of causal gesture moving through space, as with performer and instrument, or agent and sound-making apparatus.
Gravitation	The attraction towards lower or higher regions of spectral space. See 'diagonal forces.'.
Holistic space	An analytical stance, realised by mentally amalgamating an array of spatial forms into a unified spatial view.
Immersive space	The filling of spectral and perspectival space in circumspace so that the listener feels immersed in the image.
Lateral space	The extension of panoramic space towards the rear of the listener.
Levitation	A relatively effortless ascent in spectral space.
Mechanised space	A source-bonded space produced by sound-emitting machines, mechanisms and technologically based systems, independently of human activity. Can participate in enacted spaces.
Mediatic space	An amalgam of spaces associated with communications and mass media, creating an image of spaces, places, distances, events, etc.
Mediatised performed space	A type of arena space within which the transmission of gestural/ensemble space is mediated by technology (microphone, loudspeakers, video) in order to preserve the intimacy of gestural space.
Microphone space	A performed gestural or utterance space where intimacy of the image is magnified. Used for creating proximate and intimate spaces in the acousmatic image.
Nested space	In perspectival space – the embracing of one space within another.
Ouverture/enclosure	Typical relational processes in perspectival space, concerned with the opening up of spatial view by ceasing proximate activity, and the (en)closing of space through introducing proximate activity which masks or cuts off any distal image.

	Associated with approach/recession.
Panoramic space	The breadth of prospective space extending to the limits of the listener's peripheral view.
Performed space	Spaces produced by intentional sound-making, as in musical performance. See 'gestural space', 'ensemble space', and 'arena space'.
Perspectival space	The relations of spatial position, movement and scale among spectromorphologies, viewed from the listener's vantage point.
Prospective space	In perspectival space – the frontal image, which extends laterally to create panoramic space.
Proximate space	The area of perspectival space closest to the listener's vantage point in a particular listening context.
Signal space	A type of behavioural space produced by the signal calls of the participants, either to communicate with each other, or to communicate their presence to other inhabitants. See also 'behavioural space'.
Source-bonded space	The spatial zone and mental image produced by, or inferred from, a sounding source and its cause (if there is one). The space carries with it an image of the activity that produces it.
Space-form	An approach to musical form, and its analysis, which privileges space as the primary articulator. Time acts in the service of space.
Spatial form	A smaller or larger spectromorphology or musical passage (but not necessarily an identifiable 'unit') that may be analysed according to its various perceived spatial attributes. Space-form is made up of spatial forms.
Spectral space	The impression of space and spaciousness produced by occupancy of, and motion within, the range of audible frequencies.
Tonal pitch space	The subdivision of spectral space into incremental steps that are deployed in intervallic combinations – a sub-category of spectral space.
Transmodal perception	The interaction and interdependence of various sense modalities.
Utterance space	A space produced by vocal sound. This may be an intimate, personal, or social space, and in communicational contexts can also be regarded as a behavioural space. Combines with 'agential space' to create 'enacted space'.
Vantage point	The position from where the listener views perspectival space and perceives and receives the acousmatic image. The vantage point may be fixed, variable or peripatetic.
Vectorial space	The space traversed by the trajectory of a sound, whether beyond or around the listener, or crossing through egocentric space.

Table 9.1: Smalley's Space Form Glossary

Spatial Artmaking Strategies

Part two of the chapter concerns the practice of *composing with space.* While it is maintained that the psychological underpinning of spatial

forms is highly relevant in an archetypal sense to musical composition and sound appreciation in general, the domain of multichannel acousmatic music provides a direct way to apply the language. The practical techniques presented in this section of the chapter should not be seen as exclusive, as multichannel loudspeaker arrays may of course be combined with any artform. Spatial artmaking strategies differ from spatial observations made in the natural world, in that the spatial constructs presented need not necessarily comply with laws of the natural universe. For example, *motion pathways* may be wildly exaggerated, and large spaces may be nested in smaller ones. Gary Kendall explains this creative approach as follows:

> In electroacoustic music, one of the ways in which spatial events take on artistic meaning is by violating the acoustic behaviours that we associate with image schemas. Kendall (2007) discusses how the typical acoustic behaviours associated with the schema Containment can be violated by signal processing to situate listeners inside a container where they should not be. (This happens in Denis Smalley's Empty Vessels when the listener's perspective is shifted inside of a large garden pot.) (Kendall 2010b: 232)

The SoundDome Project

From 2011 to 2018 a special project under the name *SoundDome* was carried out at the University of Auckland and in an industrial warehouse in Matangi, Waikato, New Zealand. The project was focussed on the paradigm of *composing with space*. As part of the research, the psychological effects of known spatial forms were studied, and a range of associated spatialisation techniques for the *SoundDome* were developed. The near-anechoic thirty square metre space with twenty-two discrete equidistant loudspeakers, a close pair, and subwoofer (twenty-five loudspeakers in total) is pictured in figure 9.1.

Fig. 9.1 Inside the SoundDome

A leading concept emerging from the *SoundDome* project concerns the way in which spatial forms may be *exaggerated* for artistic effect. This approach known as *cartoonification* involves not only the stripping away of unnecessary information, but also 'the exaggeration of particular information features to facilitate intended perceptions unambiguously' (Lennox, Vaughan and Myatt 2001: 5). Within the context of the *SoundDome,* cartoonification processes were often used to promote an exaggerated sense of movement in foreground sounds while three-dimensional ambisonic field recordings were most often used to promote a sense of ultra-realism in background textures. The terms *proximal and distal*, *vertical*, and *surround* will now be used to present the range of applied spatialisation techniques that emerged from the project.

Proximal and Distal Spatialisation Techniques

It is well known that the high-frequency portion of the sound field is the carrier of specific location and proximity cues. Lennox, Vaughan and Myatt (2001) claim that 'this provides a further cartoonification opportunity, because information relating to proximity, location, movement, behaviour and so on can be very successfully presented using high frequency cues alone.' (Lennox, Vaughan and Myatt 2001: 6) From a production standpoint sounds that demonstrate *comingness* are easily

manufactured through a roll-on of high frequencies. Cartoonification also takes a special role in this process. The broader the brushstrokes – that is, the more obvious the frequency filtering, the more successful the effect will tend to be. Such sounds may be used deliberately to create dramatic compositional effects. Within the context of the *SoundDome project,* they tended to be particularly successful when combined with abstract musical, and sound-based cadential structures.

In multichannel acousmatic music, *close* speakers (near the listener's head) are particularly effective in delivering intimate sounds. In the *SoundDome,* in addition to the main array, the physical placement of small, full bandwidth loudspeakers as close to the listeners' ears as possible produced excellent results, allowing both distal and proximal field sounds to be presented simultaneously. Several proximity array configurations were tested including sound-beam technologies and a three-tiered multiple-distance array. The two-tiered array (ultimately adopted) delivered the most pronounced of the proximal/distance-related effects – the ability to present full bandwidth textures on the periphery as well as intimate sounds in the close pair *without bandwidth interference.* Surprisingly, our human ability to interpret space seems to allow at least close and far zones to be attended *simultaneously* allowing a kind of doubling of spectral bandwidth associated with each respective space.

In multichannel production environments, disconnects often arise between the physical location of loudspeakers and the proximity cues contained in the sonic materials. In these situations, it is most often the *'what'* associations that override other competing *'where'* characteristics. For example, an intimate vocal recording is still heard as *intimate* irrespective of the location of the loudspeaker from which it is projected - although distortions in perceived quasi-visual size and energy often occur.

Vertical Spatialisation Techniques

The initial *SoundDome* design included matched and equidistant full bandwidth loudspeakers in the upper array; however, this was eventually deemed unnecessary as bass-heavy sounds play little or no role within this upper spatial domain. Subsequent tests have found that fewer speakers of lower specifications are satisfactory to produce effective sonic images within this upper zone. Gary Kendall provides the following comments that support these observations:

> Elevation perception is the least accurate. Directional hearing above the head is very imprecise and directional hearing below the body is almost non-existent. Importantly, the acoustic cues to elevation created by the pinnae are confounded with the spectral characteristics of the sound source. In general, the higher the spectral energy distribution of the source, the higher the perceived elevation. This is verified in listening tests with a variety of filtered signals (Bloom 1977; Roffler and Butler 1968a, 1968b; Susnik, Sodnik and Tomazic 2005). In general, high-frequency sound events appear more elevated than low-frequency events, and bright events appear more elevated than dark ones. (Kendall 2010b: 231)

Of special interest, were tests carried out comparing panophonic (flat 8) vs. floor-ceiling (vertically expanded to 16) loudspeaker arrays. A common practice within the international sonic arts community is to retain the flat-8 array, and add height through the addition of an upper array; however, it was found that shifting the first 8 loudspeakers to the floor (relying on virtual placement to produce a mid-point image) provided a noticeably improved sense of immersion, which seemed desirable given the focus on *ambisonics* at the time. This also allowed accurate location-specific placement of any sound on the periphery of the array from the position of the floor to well above head height.

The vertical domain also seems to be one that is heavily influenced by psychoacoustics. For example, the sounds of birds are likely to produce the impression of being located above head height irrespective of their physical loudspeaker location. Other sounds, such as the ticking of a clock, might be deliberately placed within the upper zone to enhance

specific psychological qualities such as the 'strong geometry' as discussed by Bachelard (1964).

Surround Spatialisation Techniques

As part of the *SoundDome* project, 1st, 2nd and 3rd-order ambisonics were rigorously tested. Research available from the international sonic arts community at that time suggested the following:

1. High order ambisonics provide the largest obtainable sweet spot – but this is at the cost of near-field information (proximity/location cues etc). Location cues on the periphery are at their most pronounced.
2. Low order environmental recordings can provide a sense of 'immersion' and more reliable near-field information - but this is at the cost of a small, sweet spot and less reliable spatial information on the periphery of the array.
3. The combination of high-order-synthesised, and low-order-recorded ambisonics provides the composer with 'the best of both worlds.'.
4. The coherence of the sonic image becomes disrupted if more than one ambisonic source is used.

(Paraphrased from Barrett 2012)

This research heavily supports the superiority of higher-order ambisonics both in terms of location-specific data and the quality of three-dimensional imaging (see Gerzon 1971 & 1980, Barrett 2012, and Wilson and Harrison 2010), yet *surprisingly,* tests carried out in the *SoundDome* suggested the opposite. Perhaps this was due to the near-anechoic quality of the acoustic space that allowed such precision wave addition (phase differences) to occur. While some of the theoretical assumptions appeared to be accurate (assumption 2 and 4), unexpectedly, the most convincing three-dimensional sound images were produced through as few as four loudspeakers in a tetrahedral array matching the multi-capsule microphone configuration (avoiding decoding altogether). Additionally, the most successful methods of synthesising foreground movement within the space involved level differences (panning) rather than phase differences. These two methods were often combined to produce convincing foreground/background three-dimensional sonic images.

An interesting subset of spatial techniques involved the spatialisation of stereo sounds within an ambisonic field (on the periphery of the array). This became problematic at times since almost all stereo sounds contain multiple proximity/spatial cues. This led to the concept of spectral splitting and physical placement of foreground/background materials. It also led to the acquisition of materials (field recordings) in stems – that is the simultaneous or asynchronous recording of multiple spatial zones within the source materials themselves. As part of this sub-project a series of *zoned* multichannel field recordings were carried out, many of which involved the creative synthesis or construction of such stems on location (in the field) – as discussed in *Chapter Eight: Environmental Connections.*

The *SoundDome* project also investigated techniques concerning the development of convincing *hyper-real* surround-sound landscapes that were being concurrently researched by other members of the international sonic arts community working with similar technology. Barry Truax outlines one such technique:

> As soundscape compositions incorporate both narrative and abstracted elements into their language, a further development is to create completely imaginary yet hyper-realistic spaces. Purely synthetic sound design for environmental sounds lags behind the synthesis techniques for instrumental and vocal sound that have been developed over the last decades, particularly in terms of the possible complexity and realism of the sound material. On the other hand, processing of recorded environmental sounds tends to suggest a level of abstraction that leads one away from realism, even if fruitfully as described above. I have discovered an interesting hybrid approach that involves convolution of different classes of environmental and synthetic sounds that I used exclusively to create my eight-channel work Chalice Well (Truax 2012: 197).

Such practices promote the possibility of creating physical-spatial realities for unreal sounds and unreal spaces (acousmatic Frankensteins) – where, as previously mentioned, plausibility is maintained, even though such spaces could not exist in the natural world.

The Sensory Sphere

The *SoundDome* was designed around the concept of a perfect sphere with a single pair of ears in the centre (the sweet spot). Interestingly, the ideologies surrounding panophonic or three-dimensional *ambisonics* appear to dominate many modern studio designs of this type. Matched and equidistant monitors, in high-order configurations (8, 16, 32, etc) tend to be the desired norm. While this approach was highly successful in terms of the *imaging* achieved in the *SoundDome*, it was completely at odds with the psychologically convincing audio sphere that was observed as being *useful* to composers and sonic artists working creatively in the space. Unexpectedly, this *sensory sphere* appeared to cross-reference directly with the visual periphery of operators sitting (without moving) in the space. Head turns also played a small part in extending this imaginary elliptical space; however, most auditory attention remained firmly fixed on the foreground – producing a forward-looking auditory image that extended well beyond the forward-facing wall of the *SoundDome* enclosure. Those interested in obtaining a clearer view of the exact shape of the *sensory sphere* might do so by identifying the boundaries of their visual periphery by moving their hands in and out of optical view. This approach to multi-channel configuration design seems to share many similarities with 'acousmonium' designs – the term taking its name from the Radio de France configuration in Paris, which employs dozens of loudspeaker pairs, near-far, wide and narrow, with each stereo pair presenting a different set of spatial and spectral characteristics that composers are encouraged to take advantage of. Gary Kendall provides further evidence supporting the multiple stereo-pair acousmonium approach. He writes:

> It has been long known that the behaviour that we associate with phantom images in front of the listener in stereo reproduction does not continue around to the sides. Phantom images are extremely fragile on the sides. During the era of quadrophonic sound, Ratliff (1974) studied the perception of phantom images to the front, side and back of a listener. His data for the front pair of loudspeakers conformed to expectations and the rear pair produced similar results, but: The side quadrants exhibit a great degree of uncertainty … Even from the sweet spot, it is difficult to achieve a compact

> phantom image to the side. This phenomenon of spatial images shifting and spreading between loudspeakers on the side we will call image dispersion. The image dispersion effect is strongest for broadband signals and source signals made up of a mixture of events (as happens with live diffusion of stereo tracks). (Kendall 2010b: 234)

Room Acoustics

Last, but by no means least, we approach the complex topic of room acoustics. Succinctly stated, every room has a set of distinctive architectural characteristics relating to its dimensions, material construction and acoustic reflection/absorption properties that combine to present a sonic fingerprint or signature of sorts unique to that room. For example, using the formula *Frequency = Speed of Sound / (2 x Room Dimension)*, a room 6m long, 4m wide and 2.7m tall will produce fundamental standing waves at 27.7, 41.6, and 69.2Hz respectively; as well as peaks and nulls (through wave addition/subtraction) in defined patterns throughout the room corresponding to those frequencies and the harmonic partials produced by them. These lower-frequency artifacts are known as *room modes.* At higher frequencies, a range of similar complex phenomena occurs including phasing and reverberation (produced by reflective surfaces) that can radically impact on the sound imaging potentials discussed in this chapter. Viewed through this lens, every room might be considered a unique *instrument* best suited for producing certain types of sonic images and not others. In an ideal sense, a well-known room might be deliberately *activated* (through room modes, reflections, etc.) to create a series of distinctive sonic images unique to that particular room. While this approach does not support the practical need for dissemination of works in terms of transferability between spaces, it produces a corresponding ideal – that works may be composed *for* specific spaces.

Discussion and Conclusions

The narrow topic of composing with space within the context of multichannel acousmatic music will first be summarised. For a number of sonic artists, (me included) *spatial composition* offers a range of unparalleled creative potentials. The plasticity of sound without

accompanying images, the freedom to listen in imaginative ways, associations with the natural environment, combined with the ability to *compose with space* using a range of advanced spatialisation techniques, presents *multichannel acousmatic music* as a potent artistic medium. The psychological construction of the types of spaces that have been referred to as *spatial forms*, appears to be a natural process that is important to human beings as a way of making sense of the world. While facilities like the *SoundDome* provide direct access to a range of *cartoonified* spatialisation techniques that reflect the elegance and simplicity of these archetypal spatial forms, in many ways the genre of multichannel acousmatic music simply highlights what is already present in sounds. Space, it seems, is a powerful latent communicator that actively plays out in all forms of recorded or performed sound and music as well as fine arts, film, dance, theatre, and architecture. It provides an abstract means of organising and structuring sonic materials.

With aid of the wider *spatial* lens, it might therefore be possible (and useful) to reduce the essential concepts presented in this book to a series of spatial forms that might be organised in various ways for the purpose of creating highly communicative artworks. Furthermore, it is expected that the multi-space nature of this psychological space will allow the super-imposition of the multiple modes of listening identified to populate the total *space form*. A re-consideration of the main points presented in each chapter now follows with an emphasis on spatial reduction. This summary is offered as a way of highlighting the entangled nature of the eight modes of listening identified within the context of spatial composition.

A reduction of the *expressive* mode of listening produces spatial constructions of *the self* and of *elsewhere*. This simple interpretation might also be fleshed out by overlaying the *distance from me* paradigm evident in both Emmerson's (2007) and Hall's (1966) proposals. Both models suggest the emergence of very different types of spaces as the intimate expands to the immense. These examples highlight a naturally

occurring two-way blend between spatial and expressive modes of listening.

Formal musical qualities may be expressed in terms of one or more of the five image schemas: up and down, centrality, motion- linkage-causation, linearity: pathways and goals, and containment: inside and outside. These know *spatial forms* are at the heart of musical communicability and might be considered the abstract frameworks for the construction of all component and overarching structures of creative works that contain abstract musical materials (formal musical qualities). The relationship highlights a naturally occurring two-way blend between spatial and formal modes of listening.

Referential meaning in music might be reduced to the simple construction of one thing (signifier) referencing another (the sign), or the simple overlay (superimposition) of two meaning-based sounds to produce a metaphor. Fauconnier and Turner (2002) refer to these constituents as *input spaces*. Their internal components while separate, may also be seen to be mutually supportive, given that attributes within each can be seen to justify and/or highlight the other. These spaces not only support the existence of one another, but the bonds between them also create *shared spaces,* as well as developmental pathways that allow the spaces to evolve and transform over time. This method of managing spaces has been described as *conceptual blending,* where the blending of referential/spatial qualities combine to produce an understanding of one thing in terms of another. Again, natural entanglement between reference and space is apparent.

Audiovisual relationships provide yet another form of *coupled spaces* featuring hard boundaries and mutually supportive connections between audio and visual components. Homogeneity (isomorphic relationships) and heterogeneity (concomitant relationships) may be imagined as the superimposition of these two types of spaces - in the first case resolving to one schema, and in the second case with various attributes remaining separate – resulting in the activation of two different schemas that

require comparison and reconciliation. Visual-spatial blends may be further conceptualised as constituted by both audio and visual (real or imagined) materials.

Gestural associations may be regarded as spatial forms themselves occupying sonic presentation spaces from the local to the field (the event to the landscape). They may take multiple forms including the-self, other performers, AI performers, other-selves and a range of environmental agents including echoes and reverberations produced by empty spaces, environmental energies (the elements, machines etc.), imaginary agents of all types, and electrical energies. Interestingly, this category tends to highlight the numerous *species* of spatial forms that may enact and interact in agential ways within the total space. The classification is framed by the spatial-gestural blend of listening modes.

Environmental spaces and forces offer constructs that have spatial dimensions and associations with specific places drawn from life experiences. These spaces are also naturally entangled with other modes, such as visual, referential and expressive, triggering meanings and feelings through schemas held in long-term memory. Environmental space offers a counterpoint to gestural space featuring a multitude of slowly changing cyclical patterns, energies, and outpourings of natural energy that are primarily appreciated as *textural.* Such patterns may be highlighted through the concept of spatial-environmental blends of listening modes.

Finally, large-scale structures relating to the organisation of *component structures* in sonic artworks are often formed by the Gestalt principles: similarity, proximity, continuity, closure, figure-ground, and symmetry (discussed in *Chapter Ten: Structural Concepts*). All six identified principles may be represented as spatial forms; however, let us choose one example: *proximity.* The principle states: 'Playing two sounds simultaneously or close in time with one another will force a relationship to be forged between the two…'. The comment highlights the natural boundaries of the two spaces within a larger structural framework. The

framework itself (for example *arch form*) may also be seen to have spatial dimensions. The relationship highlights a naturally occurring two-way blend between spatial and structural modes of listening.

When viewed through the widest lens available we find that all and everything has spatial dimensions. Indeed, *space* seems to be a container of all fields as Bachelard (1964), Hall (1966), Munt (2002), Smalley (2007) et al suggest. A feeling, conversation, theory, family, building, mountain range, or continent may all be reduced to essential spatial forms. This *way of seeing* the world can be infectious. Our lives, it seems are populated with a glittering array of different types of spatial forms – some based on real sounds and images, others dreamed into being, some produced by known *agents* and others still arriving from unknown sources. Our daily existence provides the stage where these spatial forms enact and interact. They are the building blocks of our life experiences, and the archetypes of the language of sound-based music.

Chapter Ten: Structural Concepts

The domain of *structure* embraces a vast network of contributing concepts from micro to macro, in its widest scope including any pattern that may be represented over time. This description presents the term (structure) as a scalable concept – from the timeframes of spectromorphology to the durations of works or groups of works. Attempting to provide any semblance of a comprehensive list of structural concepts may therefore be an arbitrary task. The wide boundaries of structure might include all topics related to the development of sound that have been presented in this book to date, yet these would still only form a small part of the total range of structural concepts available. The term, nevertheless, generally connotes longer timeframes, sectioning-off structural principles that may be observed to be in common use. Within these within these narrower contexts, two such timeframes have been identified: those relating to the *development component structures*, and those concerning *structural integration*.

In *Chapter Four: Formal Musical Qualities,* various concepts relating to the development of component structures were examined. These included: the hierarchical structures of abstract music, Young's (2016) tripartite structural model - specifying moment form, morphic form, and narrative, as well as the associated principle of continuity, and the structural dimensions of image schemas. The current discussion primarily concerns *structural integration.* To aid in clarity and depth, just five interrelated concepts will be examined. These include *the referential-abstract continuum*, *auditory Gestalten*, *wholeness*, *conceptual blending*, and *peak musical experiences*. While the principle of conceptual blending has already been discussed within the context of referential meaning, it is re-examined here due to its structurally integrative capacity. The first half of the chapter offers an overview of these important principles and the roles they play in structural integration, while the second focusses on structural artmaking strategies drawn from the observations made.

As previously noted, structural forms are associated with sustained auditory attention, managed by parts of the brain (cortical pathways) that house the upper layers of mental activity (previously referred to as *contexts*, and *domains*). At this level, listeners tend to be actively engaged with accessing schemas (of various sorts) stored in long term memory, and as such, structural listening might be considered a highly subjective experience. As McAdams and Bigand explain:

> As McAdams and Bigand have suggested: 'form as experienced depends partly on the mind of the listener and partly on the structure presented to that mind' (1989: 181) 59 (quoted in Young 2016: 61)

There are, nevertheless, consistencies in the organisation of sound that have been commonly observed to participate in the process of achieving structural integration. The main principles identified will now be examined.

The Referential-Abstract Continuum

Arguably, one of the most heavily debated subjects throughout the international sonic arts community concerns the use of referential and/or abstract materials – not only for the part they play in delivering associative and emotional meanings but also as methods of structuring sonic artworks. Putting aesthetic differences aside, it is interesting to examine the structural implications of using semantic, referential and abstract materials in sonic artworks.

Katherine Norman, an advocate for the use spoken word (semantic materials) in sonic artworks states: 'I'd like to abandon musical analogies and offer oral storytelling as a new-or rather, a very old-model for performance... '(Norman 1994: 104). However, Norman, and other composers adopting similar stances are routinely criticised for their approach, due to the *cost* perceived (by some) through the use of such materials. That is, *meaning* can tend to become far less open to interpretation, thus restricting the listener's imagination. This overhead becomes apparent where abstract and semantic (and/or highly referential) materials are used together. Here, abstract musical materials

tend to *comment* on meanings that are invariably supplied and often *prescribed* through semantic/referential materials. Such is the case in much *programmatic music*. Music of this type has also been described *metaphorically* (by some) as *film music* – highlighting the way in which quasi-visual images supplied through semantic materials can dominate the listening experience.

Other types of referential materials such as environmental sounds also provide a rich resource that can supply associated structuring principles. Barry Truax provides a succinct summary of the approach that was developed as part of the World Soundscape Project (WSP), known as *soundscape composition.*

> Electroacoustic music recognises the abstract aspects of its language while acknowledging its movement towards some point of absolute abstractness, wheras soundscape composition begins in complete contextual immersion and moves towards the abstract middleground. In terms of the ballance between inner and outer complexity, phonograhy resides largely in outer complexity, with soundscape composition and some of the more abstracted forms of acousmatic music based on the interplay between the two...Two of the structural strategies I have observed in the practice of soundscape compostion (Truax 2002) are works which rely on a fixed spatial perspective (or series of such perspectives) and those which rely on moving perspectives. ... fixed perspective works imply a continuity of space... Works with movig perspectives rely on spatial transitions, usually simulated or evoked, such as the classic use of doors as transitional spaces in acousmatic music. (Truax 2008: 107)

As Truax points out, electroacoustic music (in general) tends to concentrate on 'abstract aspects of its language'. Here, decorated composers such as Francois Bayle tend to be occupied with the middle-ground tending towards the abstract. Bayle was known for working with highly abstracted *concréte* sounds as if they were physical objects undergoing a re-sculpting process. His aim of reaching this level of fluidity - where all sound-based musical parameters are at the disposal of the composer (such as those discussed by Smalley with reference to the *nine indicative fields*), is an ideology that is shared by many sonic artists, yet

his approach was routinely criticised (by some) for its lack of meaning.

To end this very brief survey of aesthetic approaches to the use of referential and/or abstract materials (with a focus on structural implications), the middle-ground stance is examined through the following passage written by Jonty Harrison (2000).

> My acousmatic journey seems to take me down a wide road. One side of the street has a sign saying 'rue Schaeffer' — here the preoccupation is with abstracting musical data from sound objects ('objets sonores'), without reference to their provenance in any anecdotal way. Activity on this side of the road springs from reduced listening and is primarily concerned with a spectromorphological exploration of sound qualities in order to discover (and, ultimately, to release) the latent structural properties of the material. This betrays a specifically 'musical' mind set, in which articulation of material and of structure dominates. The name on the other side of the road is hard to read, but I am sure that 'Schafer Street' — where natural sound is used precisely 'because' of its provenance, and where signification does not reside in the 'purely' musical (whatever that is) — cannot be more than a block away.
>
> I wander along this broad avenue in a reverie of discovery, often without a definite goal in mind, frequently stopping without warning to examine (at length!) some wonderful sound which presents itself along the way. To the outside observer, I must appear to weave a fairly drunken path, colliding with a lamp post on one side only to bounce across and graze my shins on the curb opposite. Most of the time I seem to be in a somewhat uncertain and unstable state somewhere in the middle of the road, in danger of being run down by any passing vehicle — a busload of academics perhaps, or someone driving too fast in a postmodern convertible; the groove along which that DJ is riding could intersect with any sudden lurch I might make as I stumble on, and even the acoustic ecologist on that bicycle over there could pose a threat to this unwary composer. (Harrison 2000, also quoted by Barreiro 2010: 36)

Auditory Gestalten as Structuring Principles

Gestalt principles have already been identified as contributing to both *conceptual blending* and *audiovisual* integration (discussed in earlier chapters). These powerful methods of associating objects and events with one another are drawn from decades of study within the field of

graphic arts, where the control and flow of visual information is seen as a priority. Many of these design principles have also been shown to be highly effective as methods of sound organisation. This collection of principles is known as *auditory Gestalten.*

'Gestalt' is German for *unified whole*. The first Gestalt principles were devised in the 1920's by German psychologists Max Wertheimer, Kurt Koffka and Wolfgang Kohle, who aimed to understand how humans typically gain meaningful perceptions from the chaotic stimuli around them. They identified a set of laws which address the natural compulsion to find order within disorder. According to them, the mind *informs* what the eye sees, and the ear hears by perceiving a series of individual elements as an integrated whole. The central principle surrounding Gestalt theory was neatly summarised by the Gestalt psychologist Kurt Koffka as: 'The whole is *other* than the sum of the parts.' (the colloquial phase *the whole is more than the sum of the parts* arose through mistranslation). The human eye, ear and brain perceive a unified visual or sonic image in a different way to the way individual parts of those images or sounds are perceived. This global *whole* is seen as a separate entity that is not necessarily formed by the sum of its parts. In the context of music, auditory Gestalten have wide and narrow applications. The widest scope is presented by Raúl Minsburg (2016) who highlights the way in which Gestalt principles play a part in the grouping of sounds*:* He writes:

> Leonard Meyer (1956) makes an important contribution when he decides to apply certain concepts of Gestalt psychology to musical situations. To start with, he defines texture as a *mental association of concurrent musical stimuli.* This is of particular importance for this work since he gives texture an eminently perceptive character. (Minsburg 2016: 103)

The study of auditory Gestalten has spanned at least 70 years, with several researchers advancing on the work of their predecessors. A prominent researcher active in the 80s and 90s was Albert Bregman who conducted a methodological investigation into sound perception applying Gestalt principles, with a special focus on the concept of *figure-ground*. Minsburg succinctly describes his approach:

> Bregman develops the 'auditory stream', which he defines as a perceptual unit grouping together many individual events. (Minsburg 2016: 103-104)

Bregman proposed an 'increase in number and functionality of voices' as being responsible for producing *texture* along with a simple-to-complex texture structure relating to the Gestalt principle of *figure-ground.* Interestingly, his studies do not appear to include tests relating to the concept of *proximity* – that is, where simple simultaneous or close positioning of sounds in time encourages a relationship to be forged. Others have taken up this challenge. Again, Minsburg provides comment:

> Consequently, if simultaneity is a coincidence of events and the event is developed through time, it follows that simultaneity is not strictly associated with 'verticality' of the instant (another problematic notion), but it necessarily involves a temporal elapsing as well. (Minsburg 2016: 110)

Expanding on this comment, in the sound world the concept of proximity appears to embrace both simultaneous and immediately subsequent combinations of sounds. This certainly equates with personal listening experience, where relationships between adjacent sounds are naturally sought out.

In terms of establishing a list of know auditory Gestalten, many scholars warn against a literal translation from the visual to the audio domain. Many of the numerous visual principles are based on visual impressions of *form* (for example groupings of similar objects) As music and sound unfolds over time, it is generally not open to interpretation in this way. Mark Reybrouck articulates the problem as follows:

> The extraction of meaningful constituents out of a complex acoustic input is not an arbitrary process, but is constrained by principles of grouping and dividing, as described by Gestalt psychology. Dividing, however, is an analytical process, and there is a danger of reducing the biological way of thinking to a structural description. (Reybrouck 1997: 65)

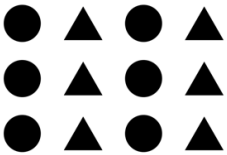
While it may not be possible to name a comprehensive list of auditory Gestalten, global principles such as similarity, proximity, continuity and closure tend to be automatically employed to help group musical elements, objects and events. Study of these principles also assists in gaining a more fine-grained conceptual understanding of the processes surrounding the appreciation of structure. Conceptually, the categories act as go-betweens providing a bridge from the domain of image schemas to the realm of conceptual blending. This relationship (between image schemas and Gestalt principles) has already been observed by several researchers – as Reybrouck explains:

> Much will depend, however, on the construction of an internal model of the external environment that can function as a trigger mechanism for Gestalt-like recognition of stimulus configurations (Sloboda 1985). For these schemas are directing our attention, in establishing correspondences between type-figures of the external stimuli and those that are stored in memory. And this, in fact, is a simple interpretation of schema theory: new structures are matched against existing ones. (Reybrouck 1997: 64)

The comment mirrors that made by McAdams and Bigand (1989) at the beginning of the chapter, that structures are formed in the minds of listeners (through the accessing of schemas); however, we should also emphasise that the *main function* of Gestalt principles is to *group* - and in the context of audio to form *relationships* between two (or more) sounds. The six principles listed in Figure 10.1 work to either suggest such *groupings*, or to *superimpose* sonic materials – in both cases leading to the integration of sonic images (from two to one image) that is characterised as being *other than the sum of its parts.* This small collection of common auditory Gestalten is presented together with paraphrased descriptions drawn from multiple sources. Visual representations of the principles have also been offered as a means of helping to establish their conceptual boundaries.

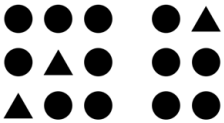
Auditory Gestalten:

Similarity:



Sounds that share sonic characteristics such as formal qualities (e.g. pitch and rhythm) or referential qualities (e.g. the sound of bells) tend to be grouped together. The further two sounds are apart over time the more similar they need to be to be recognised as a return of theme or variation or repetition of that sound. Structural forms with episodes rely on this principle to separate and group multiple themes and patterns (A, B, C. etc.) The principle relates to the containment image schema. Sounds are grouped or related to one another based on their shared sonic qualities (the 8 entangled modes identified)

Proximity:

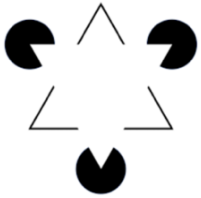


Sounds that are close together in time or auditory space are more likely to be perceived as related. This can be used to group, present, and associate sounds or component structures. Playing two unrelated sounds simultaneously or close in time with one another will force a relationship to be created between them. Metaphorical relationships tend to be sought out in the presence of semantic or referential sounds, whereas abstract sounds tend to highlight common patterns. The principle relates to the containment image schema. Sounds are grouped or related to one another based solely on their position in time.

Continuity:

Smooth transitions and flowing patterns in music, like melodic lines or rhythmic progressions, can create a sense of flow and unity. Imagine a stepwise ascending melody on one instrument, and a stepwise descending melody on another. Sustained change along a continuum (morphic structure) can also yield a sense of continuity. The principle of *movement* is at the heart of the concept, which relates to both motion-linkage-causation, and pathways and goals image schemas. Sounds are related to one another based on the position they hold within a defined structural continuum.

Closure:

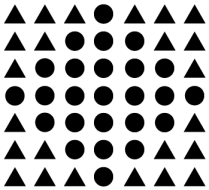


Incomplete musical phrases or patterns can be *completed* by the listener's mind. The concept calls for maximum imagination on the part of the listener. Presenting seemingly unrelated sonic images together (often using other gestalt principles) encourages in the creation of a generic image that is perceived to exist between them both. Sounds are related to one another through the concept of incomplete conceptual superimposition, where listeners are required to fill in the missing part(s).

Figure-ground:

The principle of figure-ground is based on the law of primacy of reference. Voices, or other referential materials will naturally take the foreground, along with sounds within gestural timeframe, while textural sounds - with longer or environmental phraseology will tend to take the background. For example, when distinguishing the main melodic line or a specific instrument from the background harmony or accompaniment. Sounds are related to one another through the concept of spatial superimposition, where listeners are required to construct coherent foreground-background images.

Symmetry and Order:



Symmetry occurs when two different sounds display identical patterns. For example, melodic patterns with identical intervallic structure, or two sounds demonstrating the same specific movement and spatial characteristics. The concept embraces repetitive patterns, balanced phrases, and symmetrical musical structures. Sounds are related to one another through patterns that are recognised as reflections, rotations, or translations (repeats) of previously attended patterns.

Fig. 10.1: Auditory Gestalten

The Concept of Wholeness

The concept of wholeness is presented as an important driver in the process of integrating sounds in the ways described, and in terms of the listener's need to unify works of art. It appears that this need extends to make sense of abstract patterns in terms of form (for example, pathways and goals) and the meaning of that form (for example, expressive or referential). This gives rise to the notion of local and global concerns working in parallel in the way described by the five layers of mental activity model first presented in figure 2.3. John Young presents the principle in the following comment:

> A notion of form in music is a way of thinking about how we engage with a sense of connectedness and completeness in a work. Roger Reynolds (2002) places 'wholeness' at the centre of this idea, suggesting that a composition is 'a dimensional experience that either leads a listener along a path or proposes a landscape for exploration in such a way that an arc, a trajectory of proposal, engagement, and response has been traversed by its end'. In this sense, form embraces characteristics of global shaping and its relationship to short-term clusters of events – gestures, phrases and the dynamic that exists between them. (Young 2016: 58-59)

Within the context of the current discussion, the principle of wholeness might be extended to include all the materials contained in any given work together with an inner conviction that all sounds should in some way make sense with one another. Such *certainty* forms part of the need to integrate both auditory gestalten and conceptual blends (first introduced in chapter five). These two interrelated principles will now be examined in combination.

Auditory Gestalten in Conceptual Blending

To gain a clearer understanding of where and how auditory gestalten play a role in the integration of sounds in the general manner defined by the principle of *conceptual blending* (see figure 5.3), we might first attempt to define the boundaries within which the integration of two (or more) audio streams might occur. Referring to a previous example, where the sound of a breaking wave is integrated with the sound of breathing, we

imagine simultaneous, or sequential playback of materials (reliant on the principle of proximity) to form the relationship, and we take it for granted that it would not be appropriate to attempt to present the relationship within a textural morass of sound. This is, nevertheless, an important distinction, as it defines the boundaries from within which the related principles discussed (auditory gestalten, wholeness, and conceptual bending) might play active roles. Minsburg illuminates the issue in the following comment:

> When there are three voices or streams as described above, the listener no longer practices a detailed hearing but one which is statistical, global and structural and in which a whole is perceived. What the listener privileges is the overall rather than the detailed features, though that whole is composed of individual traits. In this regard, we can bring up the definition of *textural passage* developed by Rosemary Mountain (1998) which develops along the same lines: 'When the listener's attention is drawn to the overall sonic image and the interplay of the component elements rather than on any one individual line, it becomes appropriate to describe the entire passage as *textural.'* (Minsburg 2016: 104-105)

From this description, sonic material become *textural* at the point where three or more sounds are heard simultaneously. *Textural passages* can of course feature *continuity* and *morphic form* – presenting pathways and goals that are important to the overall structure of the work*;* however, they may not be well suited to the emergence of combination images that may be seen to structurally unite disparate sounds and concepts. With this stated, it appears that auditory Gestalten (together with the principle of conceptual blending) might very well be reliant on *moment form* (or *narrative*) to present component parts. The *input spaces* of at least *similarity, proximity*, *closure* and *figure-ground* may be subject to this condition. As John Young indicates (also presented in chapter four):

> In continuing this line of thought, I am going to suggest here that a recognized *source* may have a moment function, not solely in spectro-morphological terms, but also in referential terms. (Young 2016: 69)

Returning to the example of the *wave-breath* - in cross-referencing its input spaces with the auditory Gestalten identified in figure 10.1, it is evident that it may corresponds to *several* of the principles listed, presenting the possibility that several principles may be in play at any given time. This observation suggests that the six principles identified may act as a network of integrators rather than in isolation.

Vital Relation Compression Images

Principles of auditory Gestalten are reliant on the listener identifying (or imagining) common traits between two (and sometimes three) sound objects and events. This process involves an automatic *highlighting* of common characteristics and *masking* of others in a very similar way to that which has already been described within the context of audiovisual media pairing (see *Chapter Six: Audiovisual Relationships*). Sometimes this process can create unusual, shared images, highlighting the most remote of relationships, at the cost of masking obvious features. Fauconnier and Turner refer to this process as *vital-relation compression* – noting that humans make sense of the world through a natural process of *reduction.* They write:

> One relation [between say two sounds] may be compressed into a tighter version of itself. One or more relations may be compressed into another relation. And new compressed relations may be created from scratch in the blend. (Fauconnier and Turner 2002: 310)

Fauconnier and Turner also state clearly that there is *one* overarching goal that drives the creation of a coherent vital-relation compression image – that is, to *achieve human scale*. The following comment asserts this important point:

> The most obvious human-scale situations have direct perception and action in familiar frames that are easily apprehended by human beings: An object falls, someone lifts an object, two people converse, one person goes somewhere. They typically have very few participants, direct intentionality, and immediate bodily effect and are immediately apprehended as coherent. Once blending achieves a human-scale blend, the blend also counts as human scale and so can participate in producing other human-

> scale blends, in a bootstrapping pattern that characterizes much of cultural evolution. Achievement of a human-scale blend often requires imaginative transformations of elements and structure in an integration network as they are projected to the blend. There are several subgoals worth noting, as follows: compress what is diffuse, obtain global insight, strengthen vital relations, come up with a story, and go from many to one. (Fauconnier and Turner 2002: 312)

Given this overriding concept, we may now examine some of the governing principles for compression. These are the contributors in the formation of sonic images (reflecting human scale) that are responsible for the *integration* of conceptual blends.

A common method of conceptual compression is to borrow from the human scale – that is, where one input space reflects human scale and the other does not, the second space tends to be subjected to a process of anthropomorphisation. An example of this type of vital-relations compression exhibiting the extreme malleability of the compression image is supplied by early 20th-century poet Milosz who describes the relationship between the intimacy of humans and the immensity of the world as *'two wedded creatures paradoxically united in the dialogue of their solitude'*. In this case, the poet exposes the vital-relation compression image (solitude) through which the environment is anthropomorphised.

Another method concerns compression by scaling (either up or down from the human scale) – again commonly observed where one input reflects human scale, and the other does not. This type of scaling is evident in metaphorical descriptions of music and sound. In the context of the current discussion (concerning the integration of two sounds), an example might be the integration of a highly gestural descending musical phrase (reflecting human scale) and a compressed video image of the setting sun. The result presents the environmental event in human *gestural* scale.

An important method of compression from the range identified concerns the *fabrication* of vital relations through the necessity of reconciliation. This, they state, often results where no common compression image is provided. A related method concerns instances where an overriding narrative might be adapted or extended to present a vital relation necessary for the reconciliation of two sound sources – that is, concepts supplied elsewhere (through narrative) may be used to integrate disparate sounds where no other compression image is provided. An example of narrative working in this way might be the impression of *hope* created through the emergence of a major triad in a sequence of minor chords presented alongside a spoken tale of woe. In this case, the *fabrication* of hope supplies the vital relation compression image that unites the narrative with the abstract musical materials.

Further to this last example, the use of narrative (in general) to present and create conceptual blending input spaces and vital relations compression images is worthy of special mention. Semantic materials offer unparalleled precision in the presentation of complex relationships that can, when effectively managed, supply powerful methods of structural *closure*. Referential materials of all descriptions also tend to act in quasi-narrative ways to supply the relationships required to fill the input spaces (as illustrated in figure 5.3). When referential meaning is not supplied, the necessity of achieving structural closure often falls to other modes of listening such as *formal* or *expressive*; however, conceptual blends are not easily attained through these modes alone and dimensions of dramaturgy are often sought out or fabricated by listeners in an effort to reconcile input, generic and blended spaces.

Other methods of reduction include *syncopation*, which concerns the dropping out of all but the essential elements in input spaces, and the compression of *one vital relation into another* – for example, we may attempt to combine the wave-breath compression image with Pollock's dancer-as-artist compression image leading to a further combination compressed image.

These various methods of identifying and fabricating vital relation compression images tend to form complex variants of conceptual blends including what Fauconnier and Turner (2002: 303) term 'double scope blends, multiple blends, and megablends'. To illustrate, they offer an example showing: *heat* input*, emotion* input (anger), and *body* input. To present this powerful three-way blend they simply state:

> From this we might elaborate – he was about to spontaneously combust! The image instantly assembles all three inputs while at the same time reduces the development pathway to a single moment. Vital-relation compression is one of the most powerful governing principles of conceptual integration. (Fauconnier and Turner 2002: 303)

These the highly generalised descriptions concerning methods of forming relationships between sounds presented through the various species of auditory Gestalten identified offer much food for thought concerning the effective methods available for achieving structural integration. A unifying concept presented by Fauconnier and Turner supporting the multiple species of compression methods observed concerns the human *need* to reduce and compress information for the purpose of making sense of the world. They state:

> The impulse to achieve integrated blends is an overarching principle of human cognition. Integration Principle: Achieve an integrated blend. (Fauconnier and Turner 2002: 328)

The stated principle shares many similarities with the concept of wholeness previously discussed. Fauconnier and Turner's thesis is illuminating on a number of levels in that it clearly defines: the need to integrate sonic objects and events, the required conditions for integration (non-textural), the lengths receivers go to - to integrate sounds (for example, *fabricating* compression images), and detailed descriptions of *how* we as humans make sense of one thing in terms of another.

Living with Laura (2024) – John Cousins

An example of a three-way conceptual blend features in *Living with Laura* – a multichannel acousmatic work by John Cousins. The first verse from the original 1957 song sets the scene (delivered in spoken form).

Laura is the face in the misty light
Footsteps that you hear down the hall
The laugh that floats on a summer night
That you can never quite recall
And you see Laura on a train that is passing through
Those eyes how familiar they seem
She gave your very first kiss to you
That was Laura, but she's only a dream

The work supplies several possible input spaces in the form of narrative and/or referential materials that might be seen to function as three-way blends (or multiple blends) in the way previously described. One interpretation is the combination of 1) an investigation of the lyrical content of the song (described in spoken form), 2) a musical analysis of the song (played on the piano and described in spoken form), and 3) The sound of gentle wind blowing (representing dream-like qualities referred to in the lyrics). The resulting singular *vital relation compression image* (a voice/music/wind combination image) presents a powerful metaphor that bonds with listeners' own emotions and life experiences. The composer goes further still, in reducing the compressed image to its refined essence - the slowly spoken/sung word '*L-a-u-r-a*' (containing voice/music/wind). The effect is one that achieves structural closure, uniting and cadencing the entire work.

Considering this example with Fauconnier and Turner's key principles in mind, we first find 'human scale' then attempt to convert the aspect we are trying to understand into that scale (in order of making sense of it). We then form a common reduced image (vital-relation compression

image) that represents at least two input spaces (generic space) and from this process, we often end up with images that offer a blend of input spaces as well as something *new* - something *other than* the sum of the parts. Further compression of these images is also observed through the overlay of multiple blended spaces. I would now like to suggest an expansion of this series of known structurally integrative processes to embrace the elusive concept of *peak musical experiences.*

Peak Musical Experiences

Within the bounds of structure, there may also exist the possibility of facilitating *peak musical experiences*. Before examining this relationship directly, we might first return to Nattiez's model to observe at least three species of listener: the composer, the performer and the audience, and consider what each of them may be listening *for.* A highly generalised and somewhat poetic description may present the *composer*, at the point of structural integration, attempting to resolve and unite all component structures within the work. By combining and carefully balancing, they hope to produce something extraordinary – listening for that precise sonic combination that may, in their view, deliver something profound to an audience. The *performer* seeks out this same elusive quality, often through deliberate adjustments in balance between intuitive and formal approaches to their playing strategies. When this point is attained, it is as if, in their estimation, the music achieves a higher level of quality. Drummers (for example) sometimes use the term '*in the pocket'* to define these boundaries. The *audience* also sits in wait for this moment of genesis. They too, while attending the modes of listening described above, *hope* for something more. Although there may be no conscious intention to do so, their ears often search for the point in the music where they may enter a deeply personal and significant subjective relationship with the materials. Peak musical experiences of this type are a relatively common phenomenon that most lovers of music (composers, performers, and audiences) are familiar with. In his book *Peak Music Experiences: A New Perspective on Popular Music Identity and Scenes* (2022) Ben Green explains the concept:

> Peak music experiences leave lasting impressions, informing people's future attitudes to the objects ([sounds, abstract musical materials, structural moments, artfully constructed sonic artworks], songs, artists, places) that they hold responsible for those feelings. Accordingly, these experiences offer insight into how encounters with music can affect or reorient people in enduring ways, a phenomenon to which music fans are reflexively, expectantly attuned. Peak music experiences may involve the loss or transcendence of the self in sensuous activity and collective affect, but the same experiences may also be rationalized within the narrative, purposive construction of the self. Exploring these two aspects of peak music experiences contributes to the reconciliation of two apparently competing macro-social theories of contemporary life, namely, affective sociality (e.g. Maffesoli 1996) on the one hand and reflexive individualization (e.g. Giddens 1991) on the other. (Green 2022: 9)

Some of the common qualities of peak musical experiences will now be described (paraphrased from several sources); however, it is important to note that they are not universal. They are often interpreted with such deep subjectivity that only the most rudimentary of universal characteristics may be noted. Interestingly, the range of subjective responses observed through the phenomena of *peak musical experiences* does not appear to be consistent even within individual listeners. A range of variables such as: as time of day, season, hunger levels, and what was heard *before* the listening experience can *radically* influence the emergence of this elusive quality. In extreme cases the same work (or section of work) can be deeply affecting on one day and banal on another. There is no guarantee that peak musical experiences can be repeated.

Strong Emotions

Emotional reactions to music and sound are common, but occasionally responses can be abnormally striking and even overwhelming. Research from the field of psychology reveals that the types of emotions experienced in these situations are surprisingly limited, forming just a small subset of the thirteen different emotions commonly experienced through music (as discussed in *Chapter Three: Musical Expression*). Mori and Iwanaga (2017) describe these 'peak emotions' within the context of listening to music as 'chills and tears.' They state:

> People sometimes experience a strong emotional response to artworks. Previous studies have demonstrated that the peak emotional experience of chills (goose bumps or shivers) when listening to music involves psychophysiological arousal and a rewarding effect. However, many aspects of peak emotion are still not understood. The current research takes a new perspective of peak emotional response of tears (weeping, lump in the throat). A psychophysiological experiment showed that self-reported chills increased electrodermal activity and subjective arousal whereas tears produced slow respiration during heartbeat acceleration, although both chills and tears induced pleasure and deep breathing. A song that induced chills was perceived as being both happy and sad whereas a song that induced tears was perceived as sad. A tear-eliciting song was perceived as calmer than a chill-eliciting song. These results show that tears involve pleasure from sadness and that they are psychophysiologically calming; thus, psychophysiological responses permit the distinction between chills and tears. Because tears may have a cathartic effect, the functional significance of chills and tears seems to be different. We believe that the distinction of two types of peak emotions is theoretically relevant and further study of tears would contribute to more understanding of human peak emotional response. (Mori and Iwanaga 2017: 1)

Beautiful/Functional Formal Qualities

The second characteristic that commonly contributes to peak musical experiences is the profound appreciation of formal musical qualities. While this mode of listening is also active in ordinary listening situations here, it describes a *sensual* component that often accompanies peak musical experiences. In these settings, formal musical qualities are elevated to the status of *beautiful* and *functional*. The sensation is often accompanied by feelings of deep aesthetic appreciation, exquisiteness, and awe. To frame this subtopic, I would like to draw attention to a quote by Carl Jung who said:

> "The person who speaks with primordial images speaks with a thousand tongues; they entrance and overpower, while at the same time they raise the idea they are trying to express above the occasional and the transitory into the sphere of the ever-existing… that is the secret of effective art. The creative process, in so far as we are able to follow it at all, consists in the unconscious animation of the archetype, and in the development and shaping of the image until the work is completed. The shaping of a

> primordial image is, as it were, a transition into the language of the present which makes it possible for every person to find again the deepest springs of life which would otherwise be closed to them." (Jung 1968)

To add to Jung's comments, a much-respected friend and colleague once told me 'a good work of art is like a gun – where every part has an express purpose, and when they come together in just the right way, they form an outcome that is far greater than the sum of the parts.' Here, I propose a similar explosive metaphor, but one that concerns a spark of life - the creation of something so formally intact that it reflects the beauty and complexity of the natural universe.

Insight

A third quality that commonly contributes to peak musical experiences is *insight* - a term that may connote both emotional and referential qualities. It refers to the artist's understanding of, revelation about, illumination of, or experience of something realised and expressed through the creation process; an 'aha moment', of sorts, that emerges from the organisation of materials. It is the message, knowledge, understanding, essence, realisation, epiphany, or profundity gained from an exceptional work of art – founded on the apprehension of expressive and/or referential meaning and often combined with feelings of awe. When this term is used within the context of art, it refers to a discovery, an increased knowledge, a new understanding of or about ourselves, others, and/or the world. Insight can transform, shift or expand the current views of the receiver and is made available through the materials of the work itself.

Given the divergent, ambiguous, and slippery nature of insights and the way in which they are formed, it is generally not possible to pin down specific meanings or messages within any given context; however, this does not diminish the significant roll insight has to play in the presentation and reception of music within the context of peak musical experiences. The view is also held that within the context of Western contemporary composition and sonic art, the presentation/reception of

insight is considered one of the *aims* of a high-quality work of art. Suzanne Langer uses the term with precision in the following comment, which presents her view of insight as a *combination* of expressive and referential meaning. She states:

> Yet it is perfectly clear that the 'poetic intention' which is the *raison d'être* of the work is not to give the actors [artists] self-expression, nor the audience an emotional orgy, but to *put over,* to make conceivable, a great insight into human passional nature. (Langer 1956: 182)

The concept of *epiphany* is also highly relevant, as it is clearly one that overlaps and intersects with the principle of peak musical experiences. In the following comment, Ben Green re-presents Norman Denzin's (1989, 2001) concept of an epiphany. He states:

> Epiphanies are interactional moments that leave marks on people's lives and in which personal character is manifested and made apparent. They arise out of every day, lived experience but are distinguishable from it by their form, as dramatic events with beginnings, middles and endings, and by the rupture that they represent as liminal experiences. Denzin identifies four kinds of epiphany: the major epiphany, which touches and changes every aspect of a person's life; the cumulative epiphany, which represents the culmination of a series of events; the minor or illuminative epiphany, which reveals underlying aspects of a situation of relationship; and the relived epiphany, in which an individual relives a major turning point in their life. These experiences become narrative resources with which people interpret their lives for themselves and others. Denzin notes that the meaning of experiences is given retrospectively and is never definitive; like the selves they inform, epiphanies are always unfinished productions. Further, stories of personal experience are never individual productions but are produced in and for group contexts that include cultural texts, shared histories, and criteria of truth and form. (Green 2022: 26-27)

To summarise this section, the three identified common qualities: strong emotions, beautiful/functional formal qualities, and insight proposed as characteristic of peak musical experiences may not be divisible, as it is more likely that peak musical experiences will be attended in a multimodal capacity. Nevertheless, it is interesting to observe the direct

correlation between these characteristics and the first three modes of listening identified: expressive, formal, and referential. Given this overt relationship, it is likely that other modes identified: visual, gestural, environmental, spatial, and structural may also play roles in the formation of peak musical experiences. A hallmark of the phenomenon is that they are *other than* (usually far more than) *the sum of their parts* – and with this clearly in our view, we can also deduce that highly effective forms of auditory Gestalten may be responsible for their emergence. There are also finer grained observations that support the proposed amalgamation of these two concepts (auditory Gestalten and peak musical experiences). While the eight entangled modes of listening identified can be traced to specific qualities within the sounds themselves, Gestalt relationships cannot. They are merely supported by combinations of sounds that produce unpredictable super-additive effects. Hence, the energy, (agency) of peak musical experiences tends to be reconciled by *external* qualities that are not found within the music itself. The experience is *other than the sum of the parts* and is often reconciled with notions of transcendentalism. Furthermore, the emergence of peak musical experiences can normally be traced back to specific *moments* within musical or sonic artworks. This highlights the possibility that *moment form* may be responsible for producing the experience.

Structural Artmaking Strategies

The second half of the chapter addresses the practical concerns of applying structural concepts in the artmaking process; however, before the main topic can be taken up, the creative process itself must be briefly examined to establish the boundaries of when and where the employment of structural principles may be relevant.

The Artmaking Process

There are numerous *phase models* available that may help to illuminate how the focus of activities may change from beginning to end of any given creative project. Without wishing to delve too deeply into this subtopic I would simply like to offer an overlay of my own (2005) model *Multimedia*

Production Phases and the five layers of mental activity model (Kendall 2014) referenced numerous times throughout this book.

Layer	Multimedia Production Phases
1. Sensations	Acquisition/Synthesis of Materials
2. Gist	Development of Materials
3. Locus	Development of Component Structures
4. Contexts	Structural Integration
5. Domains	Structural Refinements

Table 10.1: The five mental layers with respect to multimedia production phases

The five production phases identified may be described as divisions of a continuum that stretches from the beginning to the end of any given creative project; however, phases do not necessarily build on one another but develop all at the same time in the manner described by the five mental layers model. Nevertheless, finished concepts (which exist in the minds of composers) may be seen to reside at future positions (layers of focus) further up the figure, whereas production tasks (concerning the physical materials in any given work) reside in the present, at points further down. In this regard the model presents a top-down approach for conceptualisation, and a bottom-up approach for production. Most significantly, information may be ideally conceptualised as passing *both ways* between all production phases – with mental layers cross-checking before steering specific actions. The model highlights the necessity of all compositional concerns being active all at the same time. Where this is not observed (where composers adopt a stepwise phase model) problems can often arise due to a lack of essential qualities required to achieve structural integration. With this pointed out, it is evident that structural concepts might be considered even in the earliest phases of the formation of creative ideas.

The Genesis of Creative Ideas

One of the main challenges in the highly generalised practice of artmaking with sound is to recognise the potentials of creative ideas as they arrive. In considering the comments above, a special class of materials corresponding to all aspects of the compositional process (all

five layers simultaneously) comes into view. In my own experience as a sonic artist, this species of materials (coupled with the full gamut of creative potentials) is rarely encountered.

Applying Gestalt Principles to Conceptual Blends

A pragmatic concern of working with the principle of conceptual blending, and by extension auditory Gestalten, is that two input spaces must be filled. This often involves a process of *fabrication* (putting two sounds together) or in rare circumstances the potentials may be naturally supplied within source the materials themselves. Six methods of integration (to achieve conceptual blends) have been put forward in *figure 10.1*. These include the Gestalt principles of similarity, proximity, continuity, closure, figure-ground, and symmetry. All six species of auditory Gestalten have been associated with the human need to reduce, resolve, and understand any given audio setting. The concept of *vital-relation compression* reveals that the most satisfactory way to reconcile the combination of sonic images is through the transmutation of one or both input spaces to arrive at a generic image *reflecting human scale*. Extremely high levels of plasticity have been observed in the sound world, and even in the audiovisual world (where the companion phenomenon of *synchresis* often dominates), yet it is still surprising to observe how malleable these *vital-relation compression images* can be. As Fauconnier and Turner have pointed out, the tendency to reduce things in this way (not just in art contexts) seems to be a trait that is part to the human condition. In situations where both input spaces, and a common vital-relations compression image are discernible, conceptual blends are relatively predictable (as demonstrated in the following examples). However, where input sources are ambiguous, or compression images are not obvious, the listener is required to exercise maximum creative imagination to resolve the proposed blend. It is in these situations that auditory Gestalten tend to become unpredictable.

Examples of Auditory Gestalten Producing Conceptual Blends

We now return to and re-examine the previously discussed examples of Jackson Pollock's *No. 32,* José Halac's *The Breaking of the Scream,* and

The Author's *While the Sun Shines* for the purpose of illuminating how auditory Gestalten function in each work, and how they contribute to the spontaneous creation of conceptual blends.

Jackson Pollock's *No. 32* relies on reconciliation of a visual appreciation of abstract drip painting with an aesthetic appreciation of the human body's potential for expressive gesture through movement. To encourage this paradigm shift - from attending abstract artistic materials to the emergence of a profound concept, the artist has provided his audience with a clue - a *bridge* – an understanding of how the artwork was made, together with life-sized brushstrokes that can be recognised as the movements of Pollock's arms and body. Several Gestalt principles are arguably in play to help bridge this gap including figure-ground and continuity, but it is the principle of *similarity* working together with *closure* that arguably houses the raison d'être of the work. As figure 10.1 states:

> Sounds [as well as images and concepts] are grouped or related to one another based on their shared sonic [in this case visual] qualities (defined by the eight entangled modes identified).

In this case, the shared pattern identified relates to both brushstrokes and life-sized human gestures (the movement of limbs). These two images are resolved in a process of *closure* involving the fabrication (on the part of the viewer) of a generic image. Regarding closure, Figure 10.1 states:

> Sounds [and images and concepts] are related to one another through the concept of incomplete conceptual superimposition, where listeners are required to fill in the missing part(s).

In this case, one such interpretation creates the *artist-as-dancer* metaphor, which quickly expands to present the message: *we are all dancers in this great artwork we call life.* In the moment of discovery (or fabrication), this message is so profound that we find it hard to believe that the artist was not aware of it from the earliest creative impulses that

began the work, yet, given the unpredictable nature of Gestalt principles, it is more likely that Pollock discovered this potential (or something close to it) *during* the process of realisation. The moment of crystallisation may even have arrived by accident - through a series of experiments that he was carrying out. The example is one that moves seamlessly from expressive gestural impulses to structural artistic form, to the emergence of profound referential meaning. It ultimately can be seen to present all three peak experiences: strong emotions, beautiful and functional formal qualities, and insight.

José Halac's *The Breaking of the Scream* relies on the listener's ability to recognise and appreciate abstract patterns within gestural timeframes (the scream) and conceptually reconcile those patterns with those found in structural timeframes (the duration of the work). The discovery of this concept (deliberately employed by the composer) forces a retrospective enquiry of both the initial 60-second scream and the entire piece, adding an overwhelming level of complexity and expression to both while at the same time structurally uniting the work. Several Gestalt principles are arguably in play including *continuity, closure*, and *symmetry*. In this case, the shared pattern identified is the arch form of both the scream, and the work that is reconciled through the conceptual cross referencing of the two timeframes (a re-scaling of human gesture). Interestingly, the creation of this shared image (through *closure*) is encouraged in this instance through the principle of continuity – that is supplied by the scream but *applied* to the structural development of the work (morphic form). The *constant pressure* presented through this technique arguably requires justification, which the composer provides in the form of the discovery. An explanation of the principle of continuity is drawn from *figure 10.1*.

> Sounds [and concepts] are related to one another based on the position they hold within a defined structural continuum [in this case the rescaled arch form of the scream].

Personal correspondence with Professor Halac confirms that he became aware of the work's potential to contain an epiphany of this magnitude as early as the development of component structures phase (see *Chapter Five* for further details).

The Author's *While the Sun Shines* relies on reconciliation of two component narratives through the attainment of shared metaphorical meaning. The deliberate recontextualisation of the first narrative (concerning knots and hay bales) within the context of the second (community bonds at times of loss and grief) encourages the audience to make the discovery (the metaphor of community knots). Several Gestalt principles seem to be in play to help bridge this gap including figure-ground and closure; however, it is the principle of *proximity* that is largely responsible for spearheading the metaphorical relationship. As *figure 10.1* states:

> Sounds [and concepts] are grouped or related to one another based solely on their position in time.

In this case, a shared entailment that connects both stories through the concept of 'knots being tied' is identified both literally and metaphorically by the listener. The principle is primarily employed as a means of structurally uniting and cadencing the work. It is an effective technique of construction that it naturally forms alliances with other Gestalt principles but notably in this instance the principle of *figure-ground*. *Fig 10.1* states:

> Sounds are related to one another through the concept of spatial superimposition, where listeners are required to construct coherent foreground-background sonic images.

In this case the gathering of friends and neighbours at my parent's house (described by my father and presented in recorded form) provides the background textures, while the *re-stated* semantic materials (concerning knots being tied) occupies the foreground. The two images are reconciled both spatially (through figure-ground) and conceptually (through

closure). Of special mention is the way in which the narratives, dramaturgies, feelings, schemas, and referential meanings presented in the work are subjectively transferred to the listener, becoming mirrors of their own personal life experiences. The Gestalt technique of *closure* (in this instance) may therefore be seen to be supplied by the passions, tragedies and bonds of connection experienced by individual listeners.

Like the previous examples, we find it hard to believe that the composer was not aware of the functionality of these principles from the earliest creative impulses that began the work. However, I can personally attest that the emergence of the metaphor (produced through the principle of symmetry) arrived by accident in the latter stages of structural integration, and my understanding of the principle of closure together with other Gestalt contributors did not become clear until after the work's completion.

Auditory Gestalten and Peak Musical Experiences

There appears to be two related categories of *structural principles* that are both reliant on auditory Gestalten for their functionality: conceptual blending, and peak musical experiences. We have discussed conceptual blending at length; however, I am reluctant to make any claims concerning methods of facilitating peak musical experiences. In fact, it is my firm view that they cannot be manufactured. If they do occur, it is usually by accident. I have just two brief remarks to make on the subject.

The first comment relates to the Gestalt principle of *closure.* This is a special category, as it is reliant on the imagination of the listener to *create* appropriate *vital-relation compression images* for the purpose of integrating input spaces. This is a process over which the composer has little control, as it calls for subjective interpretation on the part of the receiver. It also calls for faith (in the listener) on the part of the composer. I believe that the emergence of peak musical experiences is far more likely under these circumstances where vital-relation compression images are not supplied.

The second comment I would like to make is that - although it may not be possible to directly manufacture peak musical experiences, it is proposed that it *is* fruitful to place sounds in various configurations with one another according to the Gestalt principles listed in anticipation of happy accidents (unforeseen reactions) occurring. These experiments may be deliberately set up following the rudiments laid out in figure 10.1.

Summary and Conclusion

In the first half of the chapter, Young's (2016) tripartite structural model specifying *moment form*, *morphic form* and *narrative* is restated. The relationship between image schemas and *auditory Gestalten* is then discussed, and six *Gestalt principles* are presented: similarity, proximity, continuity, closure, figure-ground, and symmetry. The principles are unified through their ability to group or superimpose two (or more) sounds, leading to *integration* that is characterised by a resultant sonic image that is *other than the sum of its parts.* The functionality of auditory Gestalten is then examined within the context of the principle of *conceptual blending* where they are seen to act as a *network of integrators*. The human need to integrate is extended through the principle of *achieving human scale* through *vital-relation compression images* (shared qualities between input spaces) and a range of methods is then examined including the *fabrication* of common integrating images. The topic of *peak musical experiences* is then addressed, where three constituent qualities of peak experience are described: strong emotions, beautiful/functional formal qualities and insight. To end the section, a relationship between peak musical experiences and auditory Gestalten is proposed.

The second half of the chapter addresses practical concerns of applying structural concepts in the *artmaking process*. The creative process itself is first briefly examined through overlaying the author's 2005 model *Multimedia Production Phases* with the five layers of mental activity model proposed by Kendall (2014). The superimposition reveals that production phases do not necessarily build on one another but develop all at the same time in the manner described by the five mental layers

model. This highlights the necessity of considering structural concepts throughout the artmaking process – even in the earliest phases of creative genesis. The focus of discussion then turns to practical considerations surrounding the principle of *conceptual blending*, proposing two ways in which the *generic spaces* might be populated – through either observations (supplied vital-relation compression images) or fabrication (where they are not supplied). The six identified *Gestalt principles* are then re-examined as pragmatic methods of *integration* through *vital relations compression* methods reflecting *human scale.* Distinctions in the predictability of conceptual blends is also noted. Where both input spaces and obvious compression images are provided (for example, the wave-breath image) the outcome of conceptual blends may be predicted, but where a common image is not obvious, outcomes tend to be more ambiguous. *Numerous examples* of the functionality of auditory Gestalten in the context of presenting conceptual blends are put forward. To conclude the section, the possibility of facilitating *peak musical experiences* is discussed. The view is held that *they cannot be manufactured*, however, it is also suggested that where the principle of *closure* is employed (requiring imagination on the part of the listener) the emergence of peak musical experiences becomes more likely. It is also proposed that it may be fruitful to experiment with auditory Gestalten for the purpose of facilitating *happy accidents*. In support of these final comments, I offer the following passage paraphrased by Bill Green (2022).

> The defining features of Maslow's (1962) peak experiences include feelings of internal and external wholeness and harmony, timelessness and spacelessness, complete mindfulness of the present and full use of one's capacities. These experiences are felt as valuable and lead the individual to see themselves and the world as worthwhile and meaningful, as well as to seek more peak experiences, although they cannot be commanded. They are associated with the self-actualisation that represents the apogee of a hierarchy of human needs (often popularized as a pyramid). Maslow regards as healthy the ability to 'religionise' any part of life, including aesthetic experience, and he considers aesthetic peak experiences to be a central aspect of human life with long-term consequences for people's identities and perspectives. Maslow attributes peak experiences to music

> and in later work he cites music and sex as, statistically, the two easiest ways of getting peak experiences' (Maslow 1993 [1971]: 169) – (Green 2022: 30)

It seems fitting to end this chapter, and the book, with a suggestion that a network of contributors may be responsible for the emergence of peak musical experiences - a compositional ideal that has been fiercely resistant to analysis and production. Perhaps the best method of managing this lofty ambition is simply to let it go – to have faith that it will return on its own terms. Composers and sonic artists may do their best to set their traps for profound discovery, but in the end, it will be the listeners who *fill in* the missing constituents allowing transcendent musical moments to emerge. It is in these instances that we witness the materialisation of new life, new ideas, transformations, and new beginnings – bewildered in their infancy, fragile, yet vast in consequence.

A Reduction of Fundamental Principles

This breif section offers composers and sonic artists a succinct reference guide for applying the key principles outlined in each chapter to analyses of their own creative works. The summary also specifies the theoretical maximums of two, three, four, five, six, seven, and eight-way relationships between the eight modes of listening identified - providing both the producers and listeners of sound-based music with an unprecedented range of simultaneous perspectives on the functionality of sonic materials. The foundational principle on which the model is based has been termed *sonic entanglement,* which, by definition, describes the dynamic interplay of concurrent listening modes that emerge through the creation and reception of sound-based music.

The unique characteristics of each mode of listening identified are first considered in isolation (table 11.1).

Expressive
• Emotions, feelings, motion, tension, human character, identity, religious faith, and social conditions • Expressed through seriousness of intention • Communicated primarily through nuance-level sonic characteristics and narrative (semantic/referential qualities) • 'Feeling blends' with all other modes of listening • Composer/performer as agent • Composer/performer as conduit • Synthesis of agents
Formal
• Traditional musical qualities • Melody, harmony, timbre, dynamics, texture, rhythm, and form • Component structures: moment, morphic and narrative form • The pitch-time lattice • Musical patterns as image schemas: up and down, centrality, motion-linkage-causation, containment: inside and outside, linearity: pathways and goals
Referential
• Referential meaning in music (objects, events, narrative, and dramaturgy)

- Smalley's nine indicative fields: gesture, utterance, behaviour, object/substance, energy, motion, environment, vision, and space
- Semantic qualities
- Signs and symbols as meaning
- Image schemas as meaning
- Incongruous meaning through the five layers of mental activity
- Augmentation of meaning-based qualities: expressive, formal, referential, visual, gestural, environmental, spatial, and structural
- Justification of technique: abstract and abstracted syntax methodology
- Conceptual blending as meaning: two input, generic, and blended spaces
- Aristotle's formula A is to B as X is to Y

Visual

- Real and imagined audio-visual relationships
- Audiovisual vs acousmatic modes of listening/seeing
- Sychresis, structural and associationist congruence
- Isomorphic relationships
- Concomitant relationships
- Analysis-reasignment strategies, parametric mapping solutions
- A multimodal network of audiovisual integration fields
- Superadditive (Gestaltist) effects through audiovisual integration

Gestural

- Real and imagined associations with human agents as the creators of music
- The human body as the evolutionary source of music
- Event, stage, arena, and landscape spaces
- Digital musical instrument design
- Primary and secondary feedback loops
- Gesture-to-sounding-result models
- Digital doppelgangers: isomorphic and concomitant relationships
- Gesture to texture (event to landscape) pathways

Environmental

- Environment as background (a counterpoint to foreground gesture)
- Meditational modes of listening: deep listening, soundwalks, spatial listening
- Feelings of environmental awe and reverence
- Creative practices carried out in the field: stereo and multichannel recording, site-specific installations, environmentally activated instruments, work with the human body, instrumental improvisation, loudspeaker installation, sensors and stectral transformation, found instruments, and audiovisual work

Spatial
• Music and sound reduced to a series of spatial forms • Distance from the listener: intimate, personal, social, and public zones; comingness; intimacy and immensity • Vertical position: up/down image schema, pitch space, gravitation, psychological verticality • Surround location and imaging: discrete position, three-dimensional imaging, heterotopias, envelopment • Distance techniques: proximity cues, cartoonfication, speaker position, associative distance • Elevation techniques: high frequency content, associative elevation • Location and imaging techniques: level vs phase diferences, ambisonics, foreground/background, the sensory sphere • Room acoustics
Structural
• Structural integration (bringing component structures together) • Auditory gestalten: similarity, proximity, continuity, closure, figure-ground, symetry and order • The concept of wholeness (need to integrate) • Auditory gestalten in conceptual blending, managing input space relationships • Vital relation compression images: achieving human scale (borrowing, rescaling, and fabricating), further compression of images • Peak musical experiences: strong emotions, beatiful/functional formal qualities, and insight • Production phases and the five layers of mental activity, simultaneity in structuring processes • Conceptual blending as structural integration: meaning through further compression of multiple vital relation compression images • Auditory gestalten and peak musical experiences: the principle of closure

Table 11.1: A Reduction of Concepts Relating to the Eight Modes of Listening

Two-Way Modal Blends

Chapters 3-10 present numeous examples of two-way modal blends. Each chapter examines the eight entangled modes of listening identified with the aid of specific steering concepts (for example, expression) in combination with the remaining seven modes. Through this systematic approach it has become evident that the order of these two-way

combinations is irrelevant. For example, the concept of *gestual expression* may also be examined as *expressive gesture.* In mathematical terms, these combinations (where the order dosn't matter) produce what is known as triangular numbers. A simple example from a group of 4 (say 1, 2, 3, 4), produces 1-2, 1-3, and 1-4 in the first group, 2-3, and 2-4 in the second, and only 3-4 in the third – a total of 6 combinations. In applying the same principle to the eight modes of listening identified, we arive at 28 two-way combinations from the 8 original fields. These two-way relationships are presented in table 11.2.

Two-Way Modal Combinations	
Expressive-Formal	Referential-Gestural
Expressive-Referential	Referential-Environmental
Expressive-Visual	Referential-Spatial
Expressive-Gestural	Referential-Structural
Expressive-Environmental	Visual-Gestural
Expressive-Spatial	Visual-Environmental
Expressive-Structural	Visual-Spatial
Formal-Referential	Visual-Structural
Formal-Visual	Gestural-Environmental
Formal-Gestural	Gestural-Spatial
Formal-Environmental	Gestural-Structural
Formal-Spatial	Environmental-Spatial
Formal-Structural	Environmental-Structural
Referential-Visual	Spatial-Structural

Table 11.2: All Possible Two-Way Modal Combinations (order irrelevant)

Multimodal Blends

In addition to two-way blends, chapters 3-10 also present several examples of multimodal blends (between three or more modes of listening). For example, the environmental mode of listening has been discussed as being naturally entangled with at least its spatial and visual counterparts. As the chapters have progressed, up to five-way multimodal combinations have been naturally observed (for example, energy and motion pathways). To arrive at the total number of theoretical combinations we might extend the mathematical principle of

triangular numbers to consider all 2, 3, 4, 5, 6, 7, and 8-way combinations, together with the original 8 modes of listening. This produced the following set of triangular numbers (table 11.3).

All Modal Combinations
• 8 isolated modes of listening • 28 groups of two modes • 56 groups of three modes • 70 groups of four modes • 56 groups of five modes • 28 groups of six modes • 8 groups of seven modes • 1 group of eight modes
Total: 255 unique groups

Table 11.3: All Possible Modal Combinations (where order is not important)

The total number of 255 represents the number of different possible perspectives (through modal combination) that listeners might experience when attending the same sound. These relationships (between the eight modes of listening identified) may also be conceptualised as a network diagram. A graphical representation of all possible combinations was presented in chapter two figure 2.4, and is reprinted here for clarity. Repeating comments made in chapter two, the figure offers a useful way to conceptualise all possible connections and groupings that may occur between its contituent parts. Much can be gleaned from the model including voluntary focus on specfic modes of listening, two-way, three-way (and up to eight-way) modal combinations producing signature patterns and groupings - some of which tend to be more common than others. All possibilities (in terms of modal wighting and grouping) exist at the same time, with focus and attention freely shifting between groups and subsets automatically and/or voluntarily. These shifts may be further conceptualised as focus and attention *pathways* unique to each individual listener, or as *signature listening strategies* that are often replayed during repeated listening experiences.

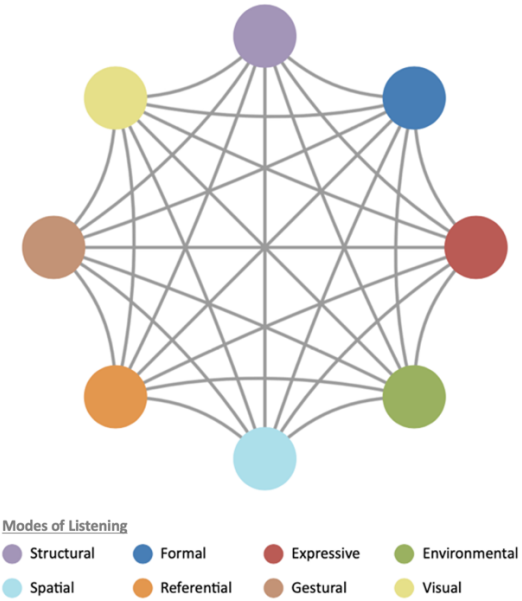


Figure 2.4 (reprinted): Eight Entangled Modes of Listening

Multilayered Multimodal Blends

The possibility of receiving incongruous messages from sonic or audiovisual materials has also been discussed in various chapters. Such incommensurate impressions (built on various groupings of listening modes) have been shown to exist simultaneously due to the five layers of mental activity: sensations, gist, locus, contexts, and domains. Differences in focus and attention (often corresponding to voluntary and automatic processes respectively) have also been noted, along with the concept of crosstalk between different processing areaas within the brain. Examples of incongruity are plentiful. A three-way blend of listening modes at the layers of sensations/gist might employ *early pathways* to embody various expressive, formal and spatial qualities, while a two-way blend at the layers of locus/contexts might employ

cortical pathways to highlight referential and structural characteristics. Such impressions may be incommensurate – that is, aspects of the two types of messages attended may not make sense when attended together. The two *images* are nevertheless commonly contracted (ignoring differences) to form the total impression of the sound and/or image under scrutiny. The concept of *incongruous layers of mental activity* allows for a total number of multimodal relationships far greater than 255 – a theoretical maximum of 1275 combinations. The significance of this final number is that it represents the number of different perspectives (through modal combination) that listeners may glean from a single listening experience.

Detailed case studies demonstrating the model's potential role in supporting compositional processes will be presented in an accompanying podcast series (available from mid-2027) and on the author's website (www.eartotheearth.nz).

References

[Adorno 2002] Theodor Adorno; *Music; Language and Composition in Essays on Music.* (Leppert, R. ed.) University of California Press, Ltd. London, England, 2002

[Atkinson 2007] Simon Atkinson, Interpretation and Musical Signification in Acousmatic Listening in *Organised Sound* 12(2): 113–22. 2007

[Austin & Smalley 2000] Larry Austin; Denis Smalley *Sound Diffusion in Composition and Performance: An Interview with Denis Smalley* Computer Music Journal, Vol. 24, No. 2. (Summer 2000) pp 10-21

[Barreiro 2010] Daniell Barreiro; Sonic Image and Acousmatic Listening in *Organised Sound* 15(1): 35–42 & Cambridge University Press, 2010.

[Barett 2012] Natasha Barrett, *The perception, evaluation and creative application of high order ambisonics in contemporary music practice* 2012. Accessed 22:08:25 from http://articles.ircam.fr/textes/Barrett12a/index.pdf

[Basanta 2017] Adam Basanta; Shades of Synchresis: A Proposed Framework for the Classification of Audiovisual Relations in Sound and Light Media Installations. eContact! 19(2) 2017 https://econtact.ca/19_2/basanta_synchresis.html (accessed 15 April 2025).

[Bachelard 1964] Gaston Bachelard *The Poetics of Space* The Orion Press, Inc, 1964. Republished by Beacon Press, Boston, Massachusetts 1994

[Bachelard 1983] Gaston Bachelard *Water and Dreams: An essay on the imagination of matter.* (Ed. Edith R Farrell) originally published in France 1942, published in English by The Pegasus Foundation Dallas 1983

[Batchelor 2015] Peter Batchelor, Acousmatic Approaches to the Construction of Image and Space in Sound Art, *Organised Sound* 20(2): 148–159 © Cambridge University Press, 2015.

[Bharucha Curtis 2008] Jamshed J. Bharucha and Meagan E. Curtis, Affective Spectra, Synchronization, and Motion: Aspects of the Emotional Response to Music in *Behavioral and Brain Sciences* 31(5) 2008.

[Bregman 1993] Albert Bregman, Auditory Scene Analysis: Hearing in Complex Environments, in S. E. McAdams and E. Bigand (Eds.) *Thinking in Sound: The Cognitive Psychology of Human Audition.* London: Oxford University Press, pp. 10-36 1993

[Bronowski 2011] Jacob Bronowski, *Science and Human Values* Faber & Faber London 2011

[Cage 1973] John Cage, *Silence: lectures and writings* Wesleyan University Press 1973

[Chion 1994], Michel Chion, *Audio-Vision: Sound On Screen* (Claudia Gorbman Ed.) Columbia University Press, New York 1994

[Clarke and Davidson, 1998] E. Clarke and J. Davidson. The Body in Performance. In W. Thomas (ed.) *Composition, Performance, Reception*. pp74-92. Ashgate Publishing Limited 1998

[Clendinning and West Marvin 2021] Jane Piper Clendinning and Elizabeth West Marvin, *The Musician's Guide to Theory and Analysis 4th ed.,* W. W. Norton & Company. 2021

[Cohen 2001] Annabel J. Cohen, Music as a source of emotion in film. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 249–272). Oxford University Press 2001

[Cook 1998] Nicholas Cook, *Analysing Musical Multimedia*, Oxford University Press Inc, New York, 1998

[Coulter 2005] John Coulter, Multimedia Composition As Research. *Proceedings of the 2005 Electroacoustic Music Studies Network International Conference:* Montreal, Quebec, Canada, 2005 Accessed online at http://www.ems-network.org/, Oct 2009.

[Coulter 2007] John Coulter, The Language of Electroacoustic Music with Moving Images. Proceedings of the Electroacoustic Music Studies Network International Conference Series: EMS07 12-15 June 2007, De Montfort University, Leicester, UK. Accessed online at http://www.ems-network.org/, Oct 2009.

[Coulter 2010] John Coulter, Electroacoustic Music with Moving Images:
The Art of Media Pairing in *Organised Sound* 15(1): 26–34 2010

[Coulter 2016] John Coulter, While the Sun Shines in *Metamorphoses 2016: 9e Concours biennial de composition acousmatique / 9th Biennial Acousmatic Composition Competition CD* Musiques & Recherches Publisher. track 4. *18.00 min* 2016

[Cousins 2019] John Cousins, *The Necessity for Consolation: John Cousins speaks* (Robert Hoskins and Norman Meehan Ed.) Te Herenga Waka University Press 2019

[Cousins 2024] John Cousins, *Living with Laura* (2024) – multichannel acousmatic work accessed from https://www.studio174-nz.com/selected-audio.html on 01/10/2025

[Cowen 2018] Allan Cowen, How Many Different Kinds of Emotion Are There? In *Frontiers for Young Minds* sourced from https://kids.frontiersin.org/articles/10.3389/frym.2018.00015 on 22/03/2023

[Cowen 2025] Allan Cowen *Emotions Evoked by Music Across Cultures* – sourced from https://www.alancowen.com/music on 06:06:2025

[De Mey 2004] Thierry De Mey *Light Music* (2004) sourced from *https://www.youtube.com/watch?v=nc6M02k2vgI&t=227s* on 27/02/2025

[EARS website] *Electroacoustic Resources Site Project* (EARS), sourced from http://ears.huma-num.fr/f9119eb1-d4a6-4bee-956e-46d464b08905.html on 06:09:2021 and 18:02:2025

[Emmerson 1986] Simon Emmerson, The Relation of Language to Materials. In *The Language of Electroacoustic Music* (S. Emmerson (Ed.)) Basingstoke, UK: The Macmillan Press, 1986

[Emmerson 1989] Simon Emmerson, Composing Strategies and Pedagogy. *Contemporary Music Review.* Harwood Academic Publishers, Vol. 3, 133-144, 1989

[Emmerson 2005] Simon Emmerson, Seeing (or not seeing) the loudspeaker; seeing (or not seeing) the music. In Relationships Between Audition and Vision in the Creation in Electroacoustic Music. (Barriére, F. and Clozier C. Ed.) Academie Internationale de Musique Electroacoustique / Bourges, Institut International de Musique Electroacoustique de Bourges / IMEB, Bourges cedex, France. Volume VIII 85-88, 2004-2005

[Emmerson 2007] Simon Emmerson, *Living Electronic Music* Ashgate Publishing Limited, 2007

[Emmerson 2015] Simon Emmerson Listening With Machines: A shared approach in *Organised Sound* 20(1): 68–75, Cambridge University Press, 2015.

[Emmerson and Landy 2016] Simon Emmerson and Leigh Landy (ed.) *Expanding the Horizon of Electroacoustic Music Analysis* Cambridge University Press 2016

[Fauconnier and Turner 2002] Gilles Fauconnier, Mark Turner, *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities,* Basic Books, a member of the Perseus Books Group, 2002

[Frengel 2010] Mike Frengel, A Multidimensional Approach to Relationships between Live and Non-live Sound Sources in Mixed Works in *Organised Sound* 15(2): 96–106, Cambridge University Press, 2010.

[Freud 1947] Sigmund Freud *The Ego and the Id* (Fourth Ed.), The Hogarth Press London 1947

[Garro 2012] Diego Garro, From Sonic Art to Visual Music: Divergences, convergences, intersections in *Organised Sound* 17(2): 103–113 & Cambridge University Press, 2012.

[Gobeil 2005] Gilles Gobeil La Perle et L'oubli (1999-2002). On *Trilogie d'ondes.* Montreal: Empreintes Digitales, IMED-0576-DVD, 2005

[Green 2022[Ben Green, *Peak Music Experiences: A New Perspective on Popular Music Identity and Scenes* Routledge Abington Oxon 2022

[Halac 1999] José Halac, The Breaking of the Scream in *Sonic Circuits CD* the American Composers Forum (ACF) track 1 *10.49 min* 1999

[Hall 1966] Edward T. Hall *The Hidden Dimension* Anchor Books Editions USA (1990 reprint) 1966

[Harrison 2000] Jonty Harrison. About E ´ vidence mate ´ rielle. http://www. electrocd.com/en/cat/imed_0052/notices/#imed_0052–0001 accessed 31:3 2026.

[Hirst 2011] David Hirst. From Sound Shapes to Space-Form: investigating the relationships between Smalley's writings and works in *Organised Sound* 16(1): 42–53 Cambridge University Press, 2011.

[Heap 2004] Imogen Heap, *Me the Machine* (2004)

https://www.youtube.com/watch?v=6btFObRRD9k accessed 21:12:23

[Hospers 2022] John Hospers, "philosophy of art". Encyclopaedia Britannica, 4 Oct. 2022, https://www.britannica.com/topic/philosophy-of-art. Accessed 30 December 2025.

[Huntley 1970] H. E. Huntley. The Devine Proportion: A Study in Mathematical Beauty, Dover Publications Inc, New York 1970

[Johnson 1987] Mark Johnson, *The body in the mind: The bodily basis of meaning, imagination, and reason.* University of Chicago Press 1987.

[Jung 1968] C. G. Jung, The Archetypes and the Collective Unconscious (Second Ed.) Bollingen Series, Princeton University Press, Princeton, USA 1968

[Juslin and Sloboda 2013] Patrik N. Juslin & John A. Sloboda, Music and Emotion in *Psychology of Music, 3rd Edition,* (Diana Deutsch Ed.) Ch. 15, 583-645 Academic Press 2013

[Kendall 2005] Gary Kendall, Qosqo: Spirituality, Process and Structure" Proceedings of the 2005 International Computer Music Conference, Barcelona, Spain 2005

[Kendall 2010a] Gary S. Kendall, Meaning in Electroacoustic Music and the Everyday Mind in *Organised Sound* 15(1): 63–74 & Cambridge University Press, 2010.

[Kendall 2010b] Gary S. Kendall, Spatial Perception and Cognition in Multichannel Audio for Electroacoustic Music in *Organised Sound* vol. 15:3 Cambridge University Press, 2010.

[Kendall 2014] Gary S. Kendall The Feeling Blend: Feeling and emotion in electroacoustic art in *Organised Sound* 19(2): 192–202 & Cambridge University Press, 2014.

[Kendall 2016] Listening and meaning: how a model of mental layers informs electroacoustic music analysis in *Expanding the Horizon of Electroacoustic Music Analysis,* Simon Emmerson and Leigh Landy (ed.) Cambridge University Press, 2016.

[Kim 2008] Suk-Jun Kim, *Listeners and Imagination: A Quaternary Framework for Electroacoustic Music Listening and Acousmatic Reasoning*. PhD thesis, University of Florida. 2008

[Kim 2010] Suk-Jun Kim, Imaginal Listening: a quaternary framework for listening to electroacoustic music and phenomena of sound-images in *Organised Sound* 15(1): 43–53, Cambridge University Press, 2010.
[Kröpfl 2005] Francisco Kröpfl, Integrating sound and visual image as artform. In Relationships Between Audition and Vision in the Creation in Electroacoustic Music. (Barriére, F. and Clozier C. Ed.) Academie Internationale de Musique Electroacoustique / Bourges, Institut International de Musique Electroacoustique de Bourges / IMEB, Bourges Cedex, France. Volume VIII (2004-2005), 89-90, 2005.

[Lakoff and Johnson 1980] George Lakoff and Mark Johnson *Metaphors We Live By* The University of Chicago Press 1980

[Lakoff and Johnson 1999] George Lakoff and Mark Johnson *Philosophy in the Flesh* Basic Books, a member of the Perseus Books Group, 1999.

[Laitz 2016] Steven G. Laitz, *The Complete Musician: An Integrated Approach to Theory, Analysis and Listening. 4th ed.,* Oxford University Press, 2016.

[Landy 2000], Leigh Landy, I Conduct Electricity. Video file. In *Devising Dance and Music: I.D..X idée fixe Experimental Sound and Movement Theatre*. (Ed. Landy, L. and Jamieson, E.) http://www.mti.dmu.ac.uk/~llandy/dance.html, Intermezzo: key company ingredients (web link), Part2.mov, 17:30-19:50, 2000.

[Landy 2007] Leigh Landy *Understanding the Art of Sound Organisation*. Massachusetts Institute of Technology 2007.

[Langer 1953] Suzanne K. Langer, *Feeling and Form*, London, Routledge and Kegan Paul. 1953

[Langer 1956], Susanne Langer, *Philosophy in a New Key: The Symbolism of Reason, Rite, and Art.* Harvard University Press, 1956.

[Lennox, Vaughan and Myatt 2001] Peter Lennox, John M. Vaughan and Tony Myatt, 3D Audio as and Information Environment published as part of the proceedings of the *19th International Audio Engineering Society Conference: Intelligent Audio Environments,* 2001.

[Lennox and Myatt 2007] Peter Lennox and Tony Myatt, A perceptual Approach to the composition of Meaning in Artificial Spatial Audio published as part of the proceedings of the *30th International Audio Engineering Society Conference: Intelligent Audio Environments,* 2007.

[Mathews 2005] Max Mathews, Music and video discussion in Relationships Between Audition and Vision in the Creation. In Relationships Between Audition and Vision in the Creation in Electroacoustic Music. (Barriére, F. and Clozier C. Ed.) Academie Internationale de Musique Electroacoustique / Bourges, Institut International de Musique Electroacoustique de Bourges / IMEB, Bourges Cedex, France. Volume VIII (2004-2005), 94-95, 2005.

[McAdams and Bigand 1993], Stephen McAdams and Emmanuel Bigand, *Thinking in Sound: The Cognitive Psychology of Human Audition.* Oxford University Press, 1993.

[McCormack, Hutchings, Gifford, Yee-King, Llano, and D'inverno 2020] Jon McCormack, Patrick Hutchings, Toby Gifford, Matthew Yee-King, Maria Teresa Llano, and Mark D'inverno, Design Considerations for Real-Time Collaboration with Creative Artificial Intelligence in Organised Sound 25(1): 41–52 © Cambridge University Press, 2020.

[McEvoy 2018] Clovis McEvoy, *Pillars of Introspection* (2018), accessed at https://www.clovismcevoy.art/pillars-of-introspection on 15:01:2025

[Minsburg 2016] Raúl Minsburg, The Relationship Between Form and Texture in *Expanding the Horizon of Electroacoustic Music Analysis* (Simon Emmerson and Leigh Landy ed.) Cambridge University Press 2016

[Miranda 2010] Eduardo Reck Miranda, Organised Sound, Mental Imageries and the Future of Music Technology: a neuroscience outlook in *Organised Sound* 15(1): 13–25 & Cambridge University Press, 2010.

[Molino, Underwood and Ayrey 1990] Jean Molino, J. A. Underwood, and Craig Ayrey, Musical Fact and the Semiology of Music in *Music Analysis* Vol. 9, No. 2, Wiley Hoboken, New Jersey, 1990

[Mori and Iwanaga 2017] Kazuma Mori and Makoto Iwanaga, Two types of peak emotional responses to music: The psychophysiology of chills and tears in *Scientific Reports* vol. 7 2017 sourced from https://www.nature.com/articles/srep46063 on 21/03/2023

[Munt 2002] Sally R. Munt *Framing Intelligibility, Identity, and Selfhood: A Reconsideration of Spatio-Temporal Models* Reconstruction: Studies in Contemporary Culture, 2 (3). 2002

[Nattiez 1990] Jean-Jacques Nattiez, *Music and Discourse: Toward a Semiology of*

Music (translated by Carolyn Abbate) Princeton University Press, 1990.

[Nguyen 2011] Xinh-Xo Nguyen, *The Road Home,* presented at the Australasian Computer Music Conference, University of Auckland 2011, accessed from https://vimeo.com/1066034815?share=copy#t=0 on 21:11:24

[Norman 1994] Katherine Norman, Telling Tales. In S. Emmerson (ed.) Timbre Composition in Electroacoustic Music, *Contemporary Music Review*, 10 (2), pp. 103-109, 1994

[Norman 1996] Katherine Norman, Real-World Music as Composed Listening A Poetry of Reality: Composing with Recorded Sound in *Contemporary Music Review*, vol 15 Parts 1-2, Taylor and Francis (formerly published by Gordon and Breach/Harwood) 1996

[Oliveros 2005] Pauline Oliveros *Deep Listening: A Composer's Sound Practice* iUniverse Lincoln NE, 2005.

[Packalén 2008] Elina Packalén, Music, Emotions and Truth, *Philosophy of Music Education Review,* vol. 16 no. 1 Indiana University Press, 2008.

[Paine 2003]. Garth Paine; Interactivity, where to from here? *Organised Sound: an international journal of music technology* 7 (3), 2003

[Pollock 1950] Jackson Pollock. *No. 32.* 1950 viewed at Tate Modern, London 15:07:2014

[Psychology Today Website] https://www.psychologytoday.com/nz/blog/the-pleasure-is-all-yours/202202/the-important-difference-between-emotions-and-feelings - accessed 06:06:25)

[Reybrouck 1997] Mark Reybrouck, Gestalt Concepts and Music: Limitations and Possibilities in *Music, Gestalt and Computing Studies in Cognitive and Systematic Musicology* (Marc Leman Ed), Springer Verlag: Berlin – Heidelberg 1997

[Riess Jones and Yee 1993] Mari Riess Jones and William Yee, Attending to auditory events: The role of temporal organization. In *Thinking in Sound: The Cognitive Psychology of Human Audition* (Ed. McAdams and Bigand) Oxford University Press Oxford. pp. 69-112, 1993.

[Roy 1998] Stéphane Roy, Functional and implicative analysis of Ombres Blanches in *Journal of New Music Research*, 27: 1, 165-184, 1998.

[Roy 2003] Stéphane Roy, *L'analyse des musiques électroacoustiques: Modéles et propositions.* Paris, L'Harnattan, Univers Musical, 2003.
[Schaeffer 1966] Pierre Schaeffer, In Search of a Concrete Music (English translation of Sheaffer's texts including Solfege de l'objet Sonore (1966), Translated by Christine North and John Dack, University of California Press, 2012.

[Schafer 1994] Murray Schafer *The Soundscape: our sonic environment and the tuning of the world* Destiny Books Vermont, 1994.

[Sloan School of Music 2023]. Sloan School of Music, *Left and Right Brain Function.* Sourced from https://sloanschoolofmusic.com/how-music-stimulates-left-and-right-brain-function/ on 19/05/2023

[Smalley 1986] Denis Smalley, Spectro-morphology and Structuring Processes. In *The Language of Electroacoustic Music* (S. Emmerson (Ed.)) Basingstoke, UK: The Macmillan Press, 1986.

[Smalley 1994] Denis Smalley *Defining Timbre – Refining Timbre* Contemporary Music Review Vol, 10, Part 2, pp 35-48 Harwood Academic Publishers 1994

[Smalley 1996] Denis Smalley, The Listening Imagination: Listening in the Electroacoustic Era in *Contemporary Music Review* Vol 13 Part 2, 1996.

[Smalley 1997] Denis Smalley, Spectromorphology: Explaining Sound Shapes in *Organised Sound* 2(2): 107–26, Cambridge University Press, 1997.

[Smalley 2007] Denis Smalley, Space-form and the acousmatic image in *Organised Sound* 12 (1) 35-58, Cambridge University Press, 2007.

[Snyder 2000] Bob Snyder, *Music and Memory – An Introduction,* The MIT Press, Cambridge, Massachusetts, London, 2000.

[Talsma, Doty and Waldorff 2006] Durk Talsma, Tracey J. Doty and Marty G. Waldorff, Selective Attention and Audiovisual Integration: Is Attending to Both Modalities a Prerequisite for Early Integration? Cerebral Cortex Advance Access publication Vol 17: 679-690, 2006.

[Taylor 2009] Byron Taylor, *Dark Green Religion: Nature Spirituality and the Planetary Future* University of California Press 2009

[Toussaint 2020] Godfried T. Toussaint, *The Geometry of Musical Rhythm* (second edition) Taylor and Francis Group, 2020

[Truax 1984] Barry Truax, *Acoustic Communication*. Norwood, NJ: Ablex Publishing Corporation, 1984.

[Truax 1993] Barry Truax, Acoustic Communication, in The Soundscape Newsletter, Number 05., March, 1993, https://www.sfu.ca/sonic-studio-webdav/WSP_Doc/Newsletters/Number5.pdf accessed 1:4:2026

[Truax 1994] Barry Truax, The Inner and Outer Complexity of Music, in *Perspectives of New Music*, 32(1), 176-193, 1994.

[Truax 2008] Barry Truax, Soundscape Composition as Global Music: Electroacoustic Music as Soundscape, in *Organised Sound*, 13(2), 103-109, 2008.

[Truax 2012] Barry Truax, Sound, Listening and Place: The aesthetic dilemma in *Organised Sound* 17(3): 193–201, Cambridge University Press, 2012.

[Wanderley 2000] Marcello Wanderley. *Gestural Control of Music.* Paris: IRCAM Centre Pompidou, Paris, 2000.

[Waters 2003] Simon Waters, H-H-H-H-Hybrids. In *ARiADA Texts 3 (May 2003)* at www.ariada.uea.ac.uk. Accessed from http://ears.huma-num.fr/a9bb424e-dd4c-472e-9f65-e31a080ccfe0.html on 23:03:2022

[Westerkamp 1974] Hildegard Westerkamp, Soundwalking, originally published in *Sound Heritage*, Volume III Number 4, Victoria B.C., 1974. Accessed from https://hildegardwesterkamp.ca/ on 27:03:2026

[Westerkamp 2019] Hildegard Westerkamp, The Disruptive Nature of Listening: Today, Yesterday, Tomorrow in Sound, Media, Ecology (Milena Droumeva and Randolph Jordan Ed.), Palgrave Macmillan, 2019

[Wilson and Harrison 2010] Scott Wilson and Jonty Harrison, Rethinking the Beast: Recent developments in multichannel composition at Birmingham ElectroAcoustic Sound Theatre in *Organised Sound* vol. 15:3 Cambridge University Press 2010

[Windsor 1995] Luke Windsor, *A Perceptual Approach to the Description an Analysis of Acousmatic Music*. PhD thesis, City University, 1995.

[Windsor 2000] Luke Windsor, Through and Around the Acousmatic: The Interpretation of Electroacoustic Sounds in *Music, Electronic Media and Culture*, Simon Emmerson (ed.), Aldershot: Ashgate, 2000.

[Wishart 1986] Trevor Wishart. Sound Symbols and Landscapes. In The Language of Electroacoustic Music (S. Emmerson (Ed.)) Basingstoke, UK: The Macmillan Press, 1986

[Wishart 1994] Trevor Wishart, *Audible Design* Orpheus the Pantomime Publishing Ltd, York, UK, 1994.

[Young 2016] John Young Forming Form in *Expanding the Horizon of Electroacoustic Music Analysis,* Simon Emmerson and Leigh Landy (ed.) Cambridge University Press 2016

[Zbikowski 1998] Lawrence M. Zbikowski *Metaphor and Music: Reflections from Cognitive Science* Music Theory Online A Publication of the Society for Music Theory Vol. 4.1 1998. Accessed from http://mto.societymusictheory.org/issues/mto.98.4.1/mto.98.4.1.zbikowski_frames.html

[Zembylas and Niederauer 2018] Tasos Zembylas and Martin Niederauer, *Composing Processes and Artistic Agency: Tacit Knowledge in Composing* Routledge New York, 2018.

About the Author

John Coulter is a New Zealand Composer/Sonic Artist and was previously the Head of Composition at the School of Music University of Auckland. His creative interests include performance-based contemporary music and sonic art, multichannel acousmatic music, environmentally driven music and art, instrument making, and sound therapy. John completed a Doctor of Philosophy degree at the University of Queensland in 2009 and has presented his work in several countries throughout the world including Australia, Canada, China, France, New Zealand, Portugal, Singapore, the United Kingdom, and the United States of America. Over his 30-year academic career he hosted several international conferences, received international awards for his creative work, was published in Organised Sound and Sonic Ideas, and founded and ran an arts-based charitable trust. He now works from his Collingwood-based studio and spends time exploring and learning from the surrounding natural environment.