Chapter 4

AN INTERMEZZO: SOUNDS AND SENSE

4.1 Introduction

At a time when computers can process massive amounts of data more impartially that humans, and cognitive science offers the elixir of computed intelligence, it may appear somewhat ironic that, not only do researchers continue to rely on wetware[1] for investigating the 'soundness' of computed results but there is a growing appreciation that audition is not a second cousin to vision, only to be brought into play when vision is unavailable or already overstrained (Bregman 1994: 1).

Data sonification is arguably the most abstract form of auditory display and historically, music is probably the most complex form of abstract aural communication made by humans. However, until very recently in human history, when recording technology afforded the production of disembodied sounds, all forms of music involved the ongoing production of corporeally-controlled interactions of physical, material objects to produce coherent sonic streams.

This short chapter, by way of an intermezzo to the consideration of sonification software, is an attempt to address in a more discursive way the question, if sonification software is to meet and anticipate the needs of sonifiers in the future, what sorts of problems will it be required to address?

4.2 Three ways of thinking about sound

The three principal ways of thinking about music can be found in three philosophers in ancient Greece in the fifth century BCE: Pythagoras, Aristoxinos and Aristotle.

[1] Described by Merriam-Webster Online as a human being considered especially with respect to human logical and computational capabilities. The Wikipedia entry describes it as a term used to describe the embodiment of the concepts of the physical construct known as the central nervous system (CNS) and the mental construct known as the human mind. It is a two-part abstraction drawn from the computer-related idea of hardware or software. Both Accessed 23 August 2008.

4.2.1 Numerical rationality

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

4.2.2 Empirical experience

Aristoxenos' approach was more concerned with the structure of the listening *experience*, which he explained in terms of the various modes[2]. A pupil of his musician father and the later Pythagoreans who were keen on distilling their inherited scientific knowledge from its more mystical entrapments, Aristoxinos' own writings on music are somewhat empirical, perhaps influenced by Aristotle, with whom he also studied. He maintained, in contradistinction to the Pythagoreans, that the notes of the scale could be tuned by the ear rather than ratio measurement, and formulated a theory that the soul is related to the body as harmony to the parts of a musical instrument[3]. Here, for example is his description of vocal pitch inflection such as glissandi:

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

4.2.3 Expressive power

Aristotle was interested in the ability of sound (music and poetry), to *express* states of mind and evoke these states in the soul (mind) of the listener.

> …for when we hear [music] our very soul is altered; and he who is affected either with joy or grief by the imitation of any objects, is in very nearly the same situation as if he was affected by the objects themselves; ... now it happens in the

---

2 Xenakis (1971: 183-189) has a more detailed explanation of Aristoxinos' modal thinking.

3 Vitruvius. *De architectura* (Book V Chapter IV) contains a paraphrase of an extant fragment of a treatise on metre in writings on music in attributable to Aristoxenos.

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

4.3 Disembodiment

Descartes' disembodied mind is a potent symbol of the armchair philosopher passively observing the world. "Sensing bodies" were considered as convenient mechanical devices to support thinking minds. While he mentions the three Greek approaches in his *Musicae Compendium* (1618), most of Descartes' effort was applied to the scientific study of the perception of objects as if by a disembodied mind (*res cogitans*). This resulted in his downgrading of the corporeal origins of musical sounds in favour of their more objective characteristics. In his *Meditations Métaphysiques* (1641), he explains that since objects have extension and can be put into motion they can be measured, however he is not sure whether qualities of objects like sounds and colours could be known:

> And in regard to the ideas of corporeal objects, I do not recognise in them anything so great or so excellent that they might not have possibly proceeded from myself; for if I consider them more closely, … I find that there is very little in them which I perceive clearly and distinctly. Magnitude or extension in length, breadth, or depth, I do so perceive; also figure which results from a termination of this extension, the situation which bodies of different figure preserve in relation to one another, and movement or change of situation; to which we may also add substance, duration and number. As to other things such as light, colours, sounds, scents, tastes, heat, cold and the other tactile qualities, they are thought by me with so much obscurity and confusion that I do not even know if they are true or false, i.e.whether the ideas which I form of these qualities are actually the ideas of real objects or not [or whether they only represent chimeras which cannot exist in fact] (Descartes, 1641 : VI).

This argument, that the unreliability of the senses made them not very suitable for building up reliable knowledge, was the subject of the Material Idealists, as discussed in §3.3.2. So, as empirical approaches to knowledge of the external world were gaining respectability, sounds were considered, like colours, as secondary properties of material objects (after Locke). Emphasising the development of objective physical knowledge of sound over

its psychological effects turned many musical issues such as tuning and temperament into targets for solutions based on logic and reasoning, the natural domain of the Cartesian mind. This had the consequence that investigation into the emotional and gestural influences, so apparent to Aristoxenos and Aristotle was neglected; it also suppressed their conscious import on musical discourse, restricting research in sound perception to 'brain–in–head' effects.

This restriction continued through the nineteenth century in experimental psychology laboratories of Wundt and Stumpf, as already discussed in §3.3.3.5. The underlying assumption in Helmholtz' psychoacoustical studies (1863) is that the sensation of *music* is based on the physiological sensing and information extraction mechanisms in the human ear, and the discoveries of representational Gestalts in perception seemed to emphasise that sensibility was achieved more through direct sensation–brain––connectivity[4] than through the experience and attention of embodied individuals.

4.3.1 Abstraction: symbolic & sub-symbolic modelling

The rise in the importance of the written score in defining musical 'works' is well documented (Goehr 1994; Attali 1985). Building on this notated representational perspective of music, the computer, assisted by MIDI-enabled keyboards for easy data input, afforded a more data–processing approach to such extra-spectral aspects of music such as melody and rhythm (Cope 1991).

In–keeping with the attempt of a burgeoning cognitive science to use computers to develop an 'artificial intelligence' which mimics human intelligence, it continued to emphasise the *res cogitans* (thinking thing) approach to musical cognition under the assumption that this was where the holy grail for musical sensibility lay (Balaban, Ebciöglu and Laske 1992).

The limited success of this approach was causal in developing sub-symbolic modeling techniques for connecting the symbolic approach to the spectral properties of sound (Kohonen 1995), aided by the stochastic techniques of the connectionists in accounting for the material constraints of musical instruments (Todd and Loy 1991). While closer to a more ecological

[4] That is, psychoneural connectivity.

approach than the previous techniques discussed, it still focused exclusively on a disembodied *res cogitans* approach to the perception and sensibility of musical sound.

4.4 The en*light*ened way

Allied to the reduction of music to noises–in–the–head was the growing dominance of visual terminology to describe aural phenomena. In his introduction to the seminal *Handbook for Acoustic Ecology*, Barry Truax provides a contemporary perspective on this occidental bias. Of the dominance of visual terminology in describing aural sensations, he writes:

> No field of study based on sensory experience seems to be overburdened by terminology to the same extent as that dealing with sound and hearing. … Terms such as perspective, foreground, background, colour, spectrum, shadow, focus, image, reflection, transparent, translucent … have found public familiarity in a way that is hard to imagine their sonic counterparts ever matching. … [T]his paradox reveals the tendency of our culture to trade its ears for its eyes (Truax 1978: v).

This privileging of the visual over the aural is implicitly supported by the sense-datum theory of perception discussed §3.3.4.1, because it accounts for immediate perceptual experience as *acts* of awareness of *objects* that such acts apprehend; When one *observes an object*, one "*watches over* a *thing aimed at*[5]. Sometimes this privileging functions to mask the unwarrantedly prejudicial, as for example in this first paragraph of *The Stanford Encyclopedia of Philosophy*'s entry on sound:

> Possibly, the philosophical privilege of the visible just reflects the cognitive privilege of the visible—as vision is considered to account for most of useful sensory information gathering. (Casati and Doric 2008).

4.4.1 Objectification

In reviewing the extensive body of literature used in chapter 3, the extent to which arguments about sensation and perception are constructed around, or exemplified by, the visual perception of spatially physical three-dimensional objects is overwhelming. The implication is that the arguments made about vision ostensibly apply equally to the other senses, and it is left up to the

---

[5] From L. *ob*-"over"+ *servare* "to watch, keep safe" and *ob*-"against"+ *jacere* "to throw," as in a jet. [OLED]

reader to induce them. Unfortunately, the problems that arise from such generalisations are embedded in language and culture.

Some researchers consider the use of the term *object* to signify a sound, as conceptually inadequate. In his seminal study of auditory scene analysis, for example, Bregman restricts the term to the visual domain and prefers the expression "auditory stream" as "our perceptual grouping of the parts of the neural spectrogram that go together … the perceptual unit that represents a single happening". In adopting *stream* he remarks that

> the word "sound" refers indifferently to the physical sound in the world and our mental representation of it. It is useful to reserve the word "stream" for a perceptual representational and the phrase "acoustic event" or the word "sound" for the physical cause. …
>
> The stream plays the same role in auditory mental experience as the object does in visual (Bregman 1994: 9-11).

While Bregman makes the uncontroversial phenomenal physical and representational distinction, his reference to the sound itself as a physical causal event is ambiguous, both in terms of the location and material composition of this "sound in the world", and whether the causal predicate is the physical "sound in the world" or the perceptual event "stream". Further, the tautological description of the origin of the sound as the sound source, does not resolve the veiled ontology of the sound itself. Bregman's work is concerned with the primitive perceptual organization of sounds, and it remains to be determined whether such perceptual organisation is affected by the omission of the consideration of the ontological nascence of either physical sounds or the auditory streams themselves.

4.5 The body returns

Husserl's pupil Martin Heidegger (1889-1976) was critical of the subject/object split that pervades the Western tradition as is indicated by the root structure of Husserl and Brentano's concept of intentionality, i.e., that all consciousness is consciousness of something, nor are there objects without some consciousness beholding or being involved with them. He encompassed terms such as 'subject', 'object', 'consciousness' and 'world' into the concept of "being-in-the-world" (*In-der-Welt-sein* or *Dasein*) and crucially, distinguished between the "present-at-hand" and the "ready-to-hand". Present-at-hand knowledge roughly corresponds to posivitist knowledge; that

required for understanding and navigating the environment–measurement, size, weight, shape, cause & effect etc. Ready-to-hand is characterized as:

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

In other words participatory, first-hand experience: familiarity, tacit know-how, skill, expertise, affordance, adaptability etc. Heidegger argues that our scientific theorizing of the world is secondary and derivative and he exposes an ontology that is far broader than the dualistic Cartesian framework. He stresses the primacy of the readiness-to-hand, with its own kind of knowing or relating to the world in terms of what matters to us.

It follows, from Heidegger's perspective, that human action is embodied, that human knowing is *enactive*, and participatory. Hubert Dreyfus (1972) used these ideas to criticise artificial intelligence, as discussed in §3.3.6, and to develop a learning model for physical skills.

The Hungarian scientist and philosopher, Michael Polanyi (1891-1976) proposed a type of participative realism in which personal knowledge plays a vital and inescapable role in all scientific research, indeed, in all human knowing.

> Let us therefore do something quite radical . . let us incorporate into our conception of scientific knowledge the part which we ourselves necessarily contribute in shaping such knowledge (Polanyi 1975: 28-9).

By stressing the tacit nature of participatory knowing, Polanyi is claiming that "we know more than we can tell." In this way he is emphasising knowledge that is implicit to a tasks (e.g. know-how, skill), to situations (e.g. navigation) and perspectives (e.g. beliefs), etc. He uses the term 'tacit knowledge' to refer to those things we can do without being able to explain how, with the way the 'body' enables one to find one's way through the world in the absence of explicit rules or calculative procedures. The "indwelling" nature of tacit knowledge is important in the development of the skill of reflexivity, such as needed in the sifting through and interpretation of qualitative data.

There seems little point in speculating whether or not listening will ever regain the relative importance to humans it enjoyed prior to the European En*light*enment, but there are signs of a growing recognition that the resolution of the mind/body dilemma will not be solved by dispensing with the body. In

many ways, the tradition of emphasising disembodied cognition over alternative approaches has never really been totally applicable to musical sensibility. The idea that musical involvement is based on the embodiment of movement and the bodily sensing of music has a long history, of which the traditional connection between dance and music is but a gross example.

Truslit studied the body movements of musical performers and suggested they were articulations of inner movements in the music itself. (Repp 1993). Central in Truslit's approach to musical movement are the notions of dynamics (intensity) and agogics (duration). If the music has the dynamo-agogic development corresponding to a natural movement, it will evoke the impression of this movement. He makes a distinction between rhythmic movement and the inner movement of the music. In contrast to rhythmic movement, which is related to individual parts of the body, the inner movement forms the melody via the vestibular labyrinth of the inner ear and is related to the human body as a whole. Both Nettheim (1996) and Clynes (1977) also make a connection between music and gravitational movement, based on the idea of a dynamic rhythmic flow beyond the musical surface.

4.5.1 Embodied cognition

Interestingly, the early criticism of the disembodied approach to the development of artificial intelligence came from inside cognitive science, in particular from philosophers (Varela, Thompson and Rosch 1991; Charmers 1995) who stress the phenomenological and embodied aspects of cognition from two perspectives. The first was that it neglects the active relationship between the perceiver and their environment and the second, that it does not take into account the activity of body components, such as eye movement in visual perception and head movement in auditory perception. (Gibson 1966; 1979; Lombardo 1987).

It is also interesting to remember that Husserl was not convinced that perception could eliminate the physical:

> The idea of a phenomenological psychology encompasses the whole range of tasks arising out of the experience of self and the experience of the other founded on it. But it is not yet clear whether phenomenological experience, followed through in exclusiveness and consistency, really provides us with a kind of closed-off field of being, out of which a science can grow which is exclusively

> focused on it and completely free of every thing psychophysical. … Ultimately the great difficulty rests on the way that already the self-experience of the psychologist is every where intertwined with external experience, with that of extra-psychical real things. (Husserl 1927/1971 §3)

A new movement-encompassing action-based approach to the relationship between sound and sensibility began in the 1980s (Cumming 2000). Other than empirical induction, the methodologies include approaches such as abductive inference (§A1.2.6.3) and are contributing to new perspectives on how to approach the relationship between and sensibilities (Varela, Thompson and Rosch 1991; Maturana and Varela 1987). In some ways this can be seen as a return to the Aristotelian integration of sound and sensibility through *mimesis* and related to the Kantian problems of openness and *endness* in the containment of beauty in formal structures and the empathic relationship within them through movement and action (Kant 1790).

4.5.2 Where is the body in sonification?

Irrespective of the sensation of (private, subjective) experiences that sounds produce in listeners, the sounds themselves clearly originate external to them and can be explored objectively to arrive at descriptions of the physical laws they obey. The sensation of sounds is clearly mediated through hearing physiology and this too can be explored empirically to arrive at an understanding of psychophysical principles involved. However, listeners are clearly more than Cartesian brains–in–vats and the perception of the immanent perceptual objects of music clearly involves more than the medium of transmission, that is, sound. For example, in an experimental study of gestures, subjects of various ages were able, with a high degree of accuracy, on only hearing different individual human's walking and running on various kinds of surfaces, to determine their sex (Bresin and Dahl 2003).

The broad ontological issues are dualistic ones: relationships between body and mind, self and other, objects and processes. They inject themselves into this present study as questions about the nature of the existence of perceptual objects and how they relate to the way listeners with bodies perceive patterns in sensations and integrate them, in conscious and unconscious ways, into their existing knowledge schemata[6].

---

[6] The term *schemata* is used here in the same sense as it was used by Piaget (1926).

Aristotle didn't include proprioception and vestibulation in his list of senses, which now should probably be augmented to at least seven. Peirce thought that there can be no perceptual object without a unifying factor that distinguishes it from the "play of impressions".

In a general way, how knowledge is acquired and validated is intertwined with the ontology of the knowledge seeker, the object of their search and the environmental contexts in which the activity occurs. The relatively recent empowering of the cognitive sciences with digital computing machines has considerably enlivened the mind-body debate, particularly with respect to the nature of consciousness, or at least some form of self-awareness (Partridge 1991).

While physical models produce convincing syntheses of individual instrumental sounds there seems to be no such model of the listener. D'haes (2004) suggests a gesture-based account of physical modelling is the way to proceed. In any event, one of the goals of music-related gesture research is to understand the biomechanical and psychomotor laws that characterise human movement in the context of music production and perception (Camurri and Volpe 2004, Camurri et al. 2005). If such an approach is fruitful, as we might hope, there is still some work to be done in generalising the human action link in the following chain to non-human action.

-> human action ->
-> sound ->
-> human empathetic movement ->
-> human sensibility/understanding ->
-> human action

Perhaps this will become the most demanding task of data sonification software in the future. Pragmatically, the next chapter attempts to address some of the requirements on software before such an approach can even begin.