Echoes in Plato's Cave: Ontology of Sound Objects in Computer Music and Analysis

Alan Marsden
Lancaster University, UK
A.Marsden@lancaster.ac.uk

ABSTRACT

The sonic aspects of Plato's analogy of the cave is taken as a starting point for thought experiments to investigate the objective nature of sound, and the idea of quasi-Platonic forms in music. Sounds are found to be objects in a way that sights or appearances are not, and it is only in the presence of technology that they become artificial. When recognition, control and communication about sound come into play, abstract concepts emerge, but there is no reason to give these the priority status Plato affords to forms. Similar issues arise in discussion of the ontology of musical works, where the ideas of extension and intension prove useful for clarity about the nature of musical objects. They are also useful for strategies in the development of music software. Musical concepts are not fixed but arise from complex cultural interactions with sound. Music software should aim to use abstract concepts with are *useful* rather than correct.

1. INTRODUCTION

Plato's theory of forms is often illustrated through his analogy of the cave (*The Republic*, Book VII). Plato has Socrates ask his hearer to imagine a cave in which prisoners have been chained since childhood in such a way that they can only see the cave wall in front of them and the shadows of objects carried past a fire. The fire and objects are behind and so cannot be seen, and Plato contends that the prisoners' perception of reality will be constituted of the shadows alone. Our perception of objects in the everyday world, according to the theory of forms, is similarly only perception of 'shadows' of forms which have a higher reality.

Most comment on the analogy, including Plato's own, is focused on vision and the sight of the shadows and, on release of the prisoner, vision of the objects which cast the shadows, of the light of the fire, and eventually of the objects of the world outside the cave and of the sun, the source of all light. However, Plato's analogy also makes reference to sound. He asks us to imagine an echo in the cave so that if those who cast shadows on the wall talked as they did so, their voices would sound as if they had come from the shadows.

Plato's point is that education allows us to see things more truly, a sentiment with which I suspect we would all readily concur, but also that some enlightened people see real truth while most of us dwell in delusion, which many of us might now regard as a dangerous idea.

I do not here want to pursue the morality or benefits of Plato's theory, but rather to explore what can be learned from following the sonic parts of the analogy, and developments from it, especially with respect to computer music and to music analysis. In both cases, a quasi-Platonic concept of an ideal 'musical object' which can be reflected in sound has been influential, but is somewhat problematic.

2. SOUND IN PLATO'S CAVE

We would now probably call Plato's analogy a 'thought experiment', so let us experiment further to investigate aspects of sound and reality. In the course of Plato's analogy, a freed prisoner is imagined to be led from the cave, and so able to see the fire and the objects which cast the shadows, and so come to understand that the shadows are not real after all. In our version of the thought experiment, let there be two freed prisoners (who do not communicate). One follows the course of Plato's original, who has been presumed to be male. Let the other be female. This freed prisoner does not see the fire and objects on the ascent from the cave (perhaps they are behind a curtain and she is too short to see over the top) but her new position allows her to hear the direct sound instead of the echo. Would that allow her to come to an understanding that the real objects have been paraded behind the prisoners and what has been seen, and heard, on the cave wall has been merely shadows? I do not think so. The direct sound would not be much different from the reflected sound. The perceived location of the source of the sound would be different, but there would be little to make clear that the newly perceived sound location is the correct one and the previously perceived location illusory.

2.1 Properties of Real Objects

Plato takes it as self-evident that the (male) freed prisoner, on seeing the objects in front of the fire, will understand that these are the real objects and that what were formerly perceived as objects are in fact shadows. With a

modern understanding of perceptual mechanisms, this is not self-evident. There is no reason to believe that he, who had only ever seen two-dimensional shadows, would have the capability to perceive the three-dimensional, and hence more real, nature of the object. Nevertheless, there are aspects of the sight of the real objects which he might perceive as richer—colour for example—and so more real. A reasonable general principle is that the illusory copy or shadow can lack properties of the real object.

The properties of the direct sound heard by the female freed prisoner are unlikely to be noticeably different from the previously heard reflected sound. Even if they were different (for example, because the reflecting surface significantly filtered the sound), there would be little to suggest that one sound was more real than the other. The difference is in balance rather than the loss of distinct properties. (A counter-example might be the common film-sound technique of heavily filtering one voice in a telephone conversation so as to distinguish between the (real) voice of the person in shot and the (artificial) voice of the other person. However, this only works because we are familiar with the technique. The filtered sound is hardly like anything we hear from a modern telephone!) The prisoner might perceive that the object had apparently changed location and changed its sound somewhat, but not that this was a real object and that what had been heard formerly was merely an echo.

2.2 Processes of Reproduction

The male freed prisoner, on seeing the fire, the objects and the shadow all at once, can come to understand the *process* by which he previously came to see the shadows. Such understanding necessarily entails a conception of the shadows previously presumed to be objects as now shadows of real objects previously not perceived.

The female prisoner has no such access to apprehension of the process by which the sound appeared previously to come from the shadows, but even if she did the previous sound would probably not cease to be real for her. When walking through an area with large sound-reflecting buildings, it is not uncommon for us to have the experience of first hearing a sound as coming from one direction and then realising that we had been hearing a reflection and that the sound really comes from elsewhere. We apprehend our mistaken belief about the object's location, but we do not apprehend a changed nature of object. For the female freed prisoner, there is no aural equivalent of the change in perception from object to shadow which there is for Plato's male prisoner.

We will consider below (Section 3) the situation when a process of sound reproduction is evident to the hearer, but even this seems not to produce a situation like that of the visual apprehension of object and shadow.

2.3 Sound objects

There seems to be no aural version of Plato's cave analogy, at least not without artificial sound-creating devices. Natural sound processes do not produce illusory objects.

Illusions can be created in sound but they do not produce illusory objects such as are supposed to be created in the minds of the prisoners by the shadows. Several well-known auditory illusions concern, as in Plato's cave, location. An example is the Scale Illusion [1] in which some notes are heard to come from the wrong location. An illusory object of sorts is created here—the non-existent scales, just as the last movement of Tchaikovsky's sixth symphony is heard to begin with a non-existent melody, composed of alternating notes from the first and second violins—but it is a different nature of object from the sounds and objects of the real world (more on this below).

Other auditory illusions concern misperception in confusing situations, such as the McGurk effect, where vision and hearing conflict, or perception of non-existent sounds continuing through interrupting noise. Illusions such as Risset's continuous glissando and sounds of ambiguous pitch might also be described as deliberately confusing: they consist of highly artificial sounds constructed in a manner to induce the auditory system to perceive sounds with properties which do not accurately reflect the physical properties.

Crucially, auditory illusions involve the misperception of sound rather than the perception of an illusory nature of sound. A shadow and a shadow-casting object are different kinds of thing, but a sound and its echo are both sounds. Even in situations when we distinctly hear the echo because it follows some time after the direct sound, we hear two sounds, not two different kinds of thing.

Sounds are objects in a way in which sights, appearances or visions are not objects. Consider another thought experiment, which perhaps corresponds to experiences you may have had. You look in a tree and believe you see a bird, perhaps an owl, sitting on a branch. On coming closer you realise that it is not a bird, but merely a twist in the branch which from your previous angle of sight looked like an owl. Your perception has changed to the degree that you now see a different object. A little further on you hear a sound coming from another tree and believe it to be a bird, a kind of crow probably, perhaps a Jay. On coming closer you see no bird but instead see an angry squirrel calling. Your perception has changed to the degree that you now perceive the squirrel to be the source of the sound rather than your previous presumption of a bird, but not to the degree of perceiving a different object. You do not hear a different sound.

The point can be argued on the basis of our use of language also. In response to your experience in the previous imagined situation you might say 'I never knew that was what a squirrel sounded like.' Note, however, that the same sentence can be used in two different circumstances. One is the situation described, where the speaker has previously heard the sound but did not realise this was the sound of a squirrel. The other is the situation where the speaker has never before heard the sound and first hears it while seeing that it comes from a squirrel. Consider the analogous sentence concerning vision: 'I never knew that was what a squirrel looked like.' This sentence can only

be used in the second kind of situation, where the speaker knows there is such a thing as a squirrel but has never before seen one. In the situation where there has been a previous misperception—perhaps the speaker has previously seen a squirrel but believed it to be a rat—the appropriate sentence would instead be 'I never knew that was a squirrel', i.e., indicating a change in perception of the object.

The phrase 'the sound of a squirrel' has a different import to the analogous 'the sight of a squirrel' or 'the appearance of a squirrel'. The first indicates something much more object-like. While Schaefferian 'reduced listening' might more strongly induce the perception of sound object (*objet sonore*) [2], it does not seem to me a necessity. Brian Kane overstates the case when he says 'A sound object is only possible when a sound no longer functions for-another as a *medium*' [3, p.18]. If this were the case, the bird/squirrel case above would yield different perceptions according to whether or not the 'other' was bird or squirrel and we would have no recognition of hearing the same sound in both situations. (One might contend that there are two modes of hearing involved in this hypothetical example, one in which sound functions as medium and one not, but this approaches rendering the definition of *objet sonore* tautological: sounds are *objets sonores* when they are heard as *objets sonores*.)

3. REAL AND ARTIFICIAL SOUND

He could not have known it, but Plato's cave has become a reality called cinema. The members of the audience are there willingly (but perhaps they are prisoners nonetheless in other senses!) and have not spent their whole lives in the cinema, but the similarity is otherwise striking: in a large dark space people view shadows on a wall (now called a screen). In place of the fire is the controlled 'fire' of a light bulb and, crucially for our purposes, the 'shadows' are not thrown by real objects but by a tiny artifice, which once was celluloid film but now in digital cinemas is usually an array of microscopic mirrors.

3.1 Reproduced Sound

Can we create an aural version of Plato's analogy in the modern era of sound reproduction technology? Suppose an unseen and not directly heard person speaks into a microphone and the sound of the voice is transmitted through loudspeakers embedded in the wall of the cave in front of the prisoners. This time also allow the female freed prisoner to see the person speaking into the microphone as well as hearing the voice directly. In contrast to the previous thought experiment, she is now likely to apprehend something of the process of sound reproduction and to understand that what had been heard formerly was, in some sense at least, artificial. Plato envisages the prisoners speaking among themselves to give words to the objects they see in the shadows. In this they would come to recognise the distinction between the speech among themselves and the speech from the loudspeakers in the cave wall.

It remains the case, though, that she will not necessarily perceive the previously heard sound as unreal. Schaeffer considers the example of hearing a recording of a galloping horse [2, p.268]. On being replayed, it is still the sound of a galloping horse, even though no horse is present. Indeed, as before, the perception of objects for the prisoners is only incorrect to the degree that the location of the speaker is misperceived. They are correct to perceive a person speaking, but only incorrect in perceiving that person as being in the cave wall rather than hidden behind them.

3.2 Artificial Sound

Now suppose that there is not a person speaking into a microphone, but a voice synthesiser transmits through the loudspeakers. In this case the freed prisoner will apprehend that what was previously perceived *was* illusory: there is no person speaking. The situation with respect to sound is now analogous to Plato's example with respect to sight. What was previously perceived to be real comes to be understood to be unreal.

Voice synthesisers can be very accurate, though, and the sound produced might be barely distinguishable from the sound of real speech. The arguments above about the object-nature of sound still apply. As sound object, the previous perception is still real. As index of someone speaking, it is unreal.

3.3 Music

Most music heard now is reproduced or artificial, and often a combination of the two. We regularly hear sounds and balances of sounds which could not be made without the use of electronic processing. I suspect that for many, now even in developing countries, it is rare to hear music which does not come from a loudspeaker or headphones. Just as for Plato's prisoners there would be a clear distinction between their own voices and the sounds from the wall, for us there is a clear distinction between everyday sounds and music: music comes from loudspeakers; everyday sound does not.

We know that Plato thought music potentially corrupting (*Republic*, book III). Would he perhaps have thought our modern electronically reproduced music most corrupting, enticing the public to remain in the cave, tickling their ears with artificial sound?

4. CONCEPTS AND FORMS

As mentioned above, Plato envisages the prisoners in the cave being able to talk to each other. He also implies that their own shadows might fall on the wall. Let us expand this again in the aural domain and allow that the prisoners are able both to communicate among themselves and to influence the sound coming from the wall. Perhaps each of them has a laptop which connects with the speakers in

the wall. (This now becomes rather like ICMC, which often takes place in cave-like rooms!)

The prisoners will come to recognise commonalities in the sounds they hear. They will come to learn to control the sounds they produce. They will develop a means of communicating with each other about the sounds. In each of these they are forming and using concepts about sound. These concepts approach much more to quasi-Platonic 'forms'—something which has an un-worldly existence, abstract and atemporal—than do the sound objects discussed previously. (Schaeffer's *objets sonores* also approach this, but only when he starts to define sounds by their abstract properties rather than by the product of reduced listening.)

4.1 Ontologies of Music

The ontology of music has been a common topic for philosophers, among whom Goodman [4], Levinson [5] and Goehr [6] are prominent. Briefly, Goodman gives an 'extensional' definition a piece of music to be the set of sound structures which conform to a particular specification of properties. Levinson gives a more 'intensional' definition as a set sound structures indicated by a particular person at a particular time. Goehr finds both problematic and argues that the concept of a musical work is historically determined, and did not come into being until about 1800.

Goehr's historical argument is compelling, and important. All of us now have always lived in a world with electronically reproduced sound—at the very least with radio, telephones and record players if not always with all the modern paraphernalia of ubiquitous digital audio. Perhaps this has influenced our conception of the objectivity of sound, and my previous arguments about the sound of squirrels might apply only to our modern age in which sound can be stored, manipulated and copied. The age before these technologies existed must have been sonically very different, but it is now unrecoverable.

My concern here with ontology differs somewhat from that of Goodman, Levinson and Goehr not only in that my historical perspective is entirely contemporary, but also in that I am concerned not just with 'musical works' but more generally with 'musical objects', which might be complete works or parts of works, or other musical components.

4.2 Extension and Intension

The notions of extension and intension are useful. The extension of a musical object is that set of things in the real world which are instances of the object (including perhaps possible and future realisations besides all actual realizations). So a note with pitch A4 is any musical sound which has fundamental frequency of 440Hz and, importantly, various other notes which are out of tune or differently but nevertheless legitimately tuned. The set of objects might be disputed or fuzzy, and might be contingent on other contextual factors.

The intension of a musical object is purely conceptual, though it may be shared. A note with pitch A4 in this conception is, roughly speaking, in our heads. We may recognise it in a sound, and we may render it in sound, but it may also be processed purely as a concept and communicated through other channels such as music notation. The definition of the intension is not by its physical properties but by the place it holds in our cognitive system of concepts.

Although this doubtless ignores important philosophical subtleties, one can equate intensions with Platonic forms (the objects casting shadows) and extensions with the shadows cast. Closer to modern sound cultures, one can equate intensions with the concepts and terms used by those who create, perceive and communicate about music, and the extensions with the sounds produced. In our last development of Plato's analogy at the beginning of this section, the intensions are indicated by the terms the prisoners use to communicate with each other, the manner in which they control the sounds, and the product of their recognition of commonalities in the sounds.

These are not new ideas, but I reiterate them because I want to make two claims:

(1) Both extensions and intensions need to be kept in mind; and
(2) Intensions (musical 'forms', 'concepts', etc.) are rarely fixed but instead can be fluid, contingent or disputable.

4.3 Keep Extension and Intension in Mind

Music theory has until quite recently generally concerned itself almost entirely with intensions, and Plato is partly to blame for this. A common and long-lasting thread of Western scholarship regards proper learning as discovering the hidden, which means being concerned with things which are not immediately sensed. Furthermore, apart from devices such as the monochord (used also by the ancient Greeks), until recently there has been little technology to allow the investigation of musical sound. Finally, music theory has been mainly concerned with the education of musicians and has defined itself by distinction from performance, which is concerned with musical sound. The result is that music theory is painfully ungrounded and, at the very least, risks making claims which do not conform to the realities of musical sound. (For discussion of an example, the concept of the 'gap-fill' melodic pattern, see [7].)

Recall that in our last analogy, intensions were considered to arise in the minds of the prisoners through recognition, control and communication. Perhaps not all of these necessarily lead to concept-formation. We can learn to control a bicycle without any conception of the mechanics involved. Could you explain to somebody else how to turn a corner on a bicycle without falling off? (If you say 'turn the handlebars in the direction you want to go', you are wrong. In fact you need to apply pressure in the *opposite* direction.) We can recognise faces but not be able to describe the features of a person which would

allow somebody else to pick them out. It is perhaps only for communication that concept-formation is essential.

So it is possible for musical processes, including music software, to operate entirely with extensions. I contend, however, that they are more adaptable, and hence more useful, if they also deal with intensions. For example, much research in Music Information Retrieval has concerned classifiers of some sort, software which determines from an audio file what class of sound or music it contains. This is commonly done by means of a machine-learning system which uses a set of training examples; in other words, the classes to be recognised are specified by (part of) their extension alone. The result is systems which can often be quite good at the classification for which they are designed, but which are otherwise useless. If a slightly different classification is required (e.g., because a new class has been introduced) the software must relearn. The system cannot be used as a basis for the design of software to perform a related task (e.g., to transform a piece of music so that it becomes a member of a different class). I do not claim that such software based on extensions only is always useless, but I suspect that usefulness is greatest when such software (a) learns continually, and (b) is embedded into real-world activities.

4.4 Musical 'Forms' are Rarely Fixed

The classic examples of Platonic forms are geometrical shapes such as the circle. These things can be given precise definition in abstract terms ('the set of points on a plane which are equidistant from a central point'). There are some musical concepts which can similarly be given precise abstract definitions (e.g., Fortean pitch class sets), but this is not true for many. One of the most important concepts in Western music is 'key', but it is difficult to define. Important factors are the use of a particular set of pitches, use of particular pitches in particular roles (e.g., tonic), and use of particular configurations of pitches (e.g., harmonic progressions), but no single combination of these by itself appears to give a solid definition of key. (For fuller discussion, see [8].)

Musical culture varies from time to time and place to place, through a complex interaction of intension and extension. The interaction is seen even in a single musician with an instrument (or computer): the musician manipulates the instrument, sound comes out, the musician hears it, and her musical ideas change. This is the excitement of music. If we focus only on sound or only on ideas, or if we believe that the ideas are fixed, we lose the excitement.

5. MUSIC COMPUTING

A significant research interest of mine has been the analysis of music by computer. Analysis can be seen as a process of going from extension to intension. We start with the sound of a piece (or its score, which is an extension from this perspective, though an intension from the perspective of the performer who aims to play the piece) and we aim to discover the forms or structures which enable us to explain the piece, or to relate it to other pieces, or to make a new piece which is similar in some respects but different in others, or to communicate about the piece to others. For some time I have been interested in Schenkerian Analysis by computer [9, 10], which explicitly aims to uncover a multi-layered musical structure underlying the notes of the piece. The outcome of this is software which is capable of finding structures in short extracts of pieces, but without great confidence that these are the right structures. The guidance from Schenker's written theory leads to multiple possible structures, without any guidance of how to choose among those many structures. It has yet to be established whether or not it is possible to learn from Schenker's examples of analysis how one should choose. The evidence from decades of students being taught Schenkerian Analysis in universities across the world is that it is possible to learn, but these students have access to far more information than just Schenker's examples.

This is an example of the lack of fixity (in the sense of lack of definition rather than changeableness) in musical forms noted above. A complex concept (Schenkerian structure) has emerged from a complex body of interaction between ideas and sounds. It lacks precise definition, but nevertheless appears to have sufficient solidity to be transmissible across time and space. Perhaps it is Schenker's writings and examples alone which provide this solidity, but I suspect not. After all, the writings and practices of other music theorists have not led to such solidity. The musical 'forms' which persist do not appear to be a random muddle. I suspect they persist because of their *usefulness*, and I consider it to be a prime desideratum for future good research in music computing to be able to define usefulness in such a way that it can be used as a criterion in judging the 'correctness' of analyses, of concepts formed in the course of software development, and in machine-learning systems.

5.1 A Useful 'Form': Tonal Profiles

I mentioned above that 'key' is a problematic concept, which means it is problematic for software which aims to determine the key of a passage of music. There still does not exist software which is guaranteed to assign the correct key to every passage of music, where 'correct' is defined as the key a majority of musical experts would assign to that passage. However, there does exist software which is often right and, more importantly, it often uses an idea which is not exactly the same as 'key' (if it were the software would always be right) but is a *useful* alternative.

Research by Shepard and Krumhansl on perception of pitch similarity [11] has led, through a number of conceptual developments, to the idea of 'tonal profile' meaning a vector of twelve values which indicate the 'fittingness' or even simple frequency of occurrence of the twelve pitch classes in a particular profile (see [12]). There is a typical profile for major keys, and another for minor

keys. Key-determining software can count the occurrences of each pitch class in a passage, and find the closest matching key profile.

The idea is not precisely a theory of key, because it does not reflect everything about that concept we find in music theory, but it is close and it is distinct and computable. Probably for these reasons, it has been extraordinarily fecund in the fields of music theory, music psychology and music computing.

5.2 A Useful 'Form' Lacking: Harmony

Music computing needs more useful ideas like tonal profiles. As an example, consider the concept of harmony. Like key, this is crucial in much of Western music. The idea of a harmonic progression underlies the basis of much improvisation in jazz and variation in classical music. It often provides the driving force in popular music also, where a piece of music can frequently be well represented as simply a melody and an accompanying sequence of chords.

In one sense the concept of harmony is straightforward: a harmony is made up of a set of pitch classes, e.g. C, E, G for C major. In practice, however, it is far from this simple. Occurrences of C-major harmony can miss out one or even two of these pitch classes, and also include occurrences of other pitch classes. Music theory distinguishes between harmony (or essential) and non-harmony (or inessential) notes, but I have yet to see a precise algorithm for making this distinction while simultaneously determining the harmony.

The situation reminds me of the situation of key-determining software before the advent of tonal profiles, and puts into my mind the probability that we need a new concept of harmony which, like tonal profiles, is distinct and computable, and which need not match everything about the music-theoretic concept of harmony but is sufficiently close to be useful.

6. CONCLUSIONS

Even without 'reduced listening', sound objects are object-like. Musical objects, when they are sound, are similarly object-like, but there are also more abstract, intensional, musical objects which are more like Platonic forms. I see no reason to consider these objects to be of greater importance or priority than sound objects, but instead the two exist in a complex cultural interaction. Abstract musical concepts are therefore subject to modifications and imprecisions arising from this complex interaction, and we should not expect music software, which must operate at some level with precise concepts, to match complex abstract musical concepts. Instead we should seek precise but useful concepts which allow productive research.

7. REFERENCES

[1] D. Deutsch, "Two-channel listening to musical scales," *J. Acoust. Soc. Amer.*, vol. 57, pp. 1156-1160, 1975.

[2] P. Schaeffer, *Traite´ des objets musicaux*. Éditions du Seuil, 1966.

[3] B. Kane, "L'Objet Sonore Maintenant: Pierre Schaeffer, sound objects and the phenomenological reduction," *Organised Sound*, vol. 12, no. 1, pp. 15–24, 2007.

[4] N. Goodman, *Languages of Art: An Approach to a Theory of Symbols*. 2nd ed. Hackett, 1976.

[5] J. Levinson, "What a Musical Work Is," *Journ. Philosoph.*, vol. 77, no. 1, pp. 5–28, 1980.

[6] L. Goehr, *The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music*, 2nd ed. OUP, 2007.

[7] A. Marsden, "Position paper: Counselling a better relationship between mathematics and musicology," *J. Math. and Music*, vol. 6, no. 2, pp. 145–153, 2012.

[8] A. Marsden, "Computers and the concept of tonality," in *Information Technology and Scholarly Disciplines*, J.T. Coppock, Ed. OUP, 1999, pp. 33–52.

[9] A. Marsden, "Schenkerian analysis by computer: a proof of concept," *J. New Music Research*, vol. 39, no. 3, pp. 269–289, 2010.

[10] A. Marsden, "Software for Schenkerian Analysis," in *Proc. Int. Comput. Music Conf.*, Huddersfield, UK, 2011, pp. 673–676.

[11] C.L. Krumhansl and R.N. Shepard, "Quantification of the Hierarchy of Tonal Functions within a Diatonic Context," *J. Experimental Psychology: Human Perception and Performance*, vol. 5, no. 4, pp. 579–594, 1979.

[12] C.L. Krumhansl, *Cognitive Foundations of Musical Pitch*. OUP, 1990.