Music and Cognitive Science

ROGER SCRUTON

It has always been controversial to make a sharp distinction between the philosophical and the psychological approaches to aesthetics; and the revolution brought about by cognitive science has led many to believe that the philosophy of art no longer controls a sovereign territory of its own. To take one case in point: recent aesthetics has addressed the problem of fiction, asking how it is that real emotions can be felt towards merely imagined events. Several philosophers have tried to solve this problem by leaning on observations in psychology – Jenefer Robinson, for example, exploring the domain of pre-conscious and non-rational responses, and Greg Currie, invoking simulation theory from the realm of cognitive science. I am not yet persuaded that either has succeeded in solving the philosophical question: but the fact that such sophisticated and well-informed philosophers should begin from studies in empirical psychology says much about how the subject of aesthetics has changed since the early days of linguistic analysis.

If philosophy has any rights in this area, however, one of them is to identify nonsense, and especially nonsense that arises when practitioners of first-order disciplines seize upon the jargon of some science in order to take an unwarranted second-order standpoint. It often seems as though cognitive science and neuroscience are now fulfilling a career need that was in a previous generation fulfilled by the postmodern theory machine. Here is Ian Cross, a respectable musicologist, taking advantage of cognitive science jargon to re-define his subject matter:

'Musics are cultural particularisations of the human capacity to form multiply-intentional representations through integrating information across different functional domains of temporally extended or sequenced human experience and behaviour, generally expressed in sound.'[1]

That is a case where we might very well call for some philosophical first aid. The unexplained plural of 'musics', the use of 'representations' to transfer our attention from the thing heard to the process

[1] In Isabelle Peretz and Robert J. Zatorre, eds., *The Cognitive Neuroscience of Music*, Oxford, OUP, 2003.

doi:10.1017/S1358246114000277 ©The Royal Institute of Philosophy and the contributors 2014

https://doi.org/10.1017/S1358246114000277 Published online by Cambridge University Press

of hearing it; the jargon, the running together of experience and behaviour and – emerging from all this like the rabbit from the hat – the extraordinary suggestion that music, whether singular or plural, is not necessarily expressed in sound. Cross is not talking about the art of sound as we know it, but about a neural process that 'integrates information' across 'different functional domains' of 'temporally extended or sequenced human experience or behaviour'. That description, in so far as it is meaningful, applies to every kind of human perception, and does nothing to distinguish music – or musics – from dinner parties and football games.

Behind Cross's nonsensical definition, however, I sense the encroachment of some real and interesting theories. First, there is the adaptation theory of the human mind, which tells us that our capacities to experience and act upon the world are the product of millennia of adaptation. Secondly there is the computational theory of the brain, according to which the brain is an information processing device that acts like a digital computer, to transform input to output through the recursive operation of quasi-syntactical algorithms. Thirdly there is the modular theory, which assigns specific 'domains' to independent capacities that evolved in order to 'process' them. All three theories have a following among philosophers, the second and the third featuring in the philosophy of Fodor as the basis for a naturalistic understanding of the mind.[2] And there are ways in which all of them might have an impact on the philosophy of music.

The adaptation theory will not concern me in what follows. Whether we follow Steven Pinker in considering music to be 'evolutionary cheese-cake', whose attractions are a by-product of other and more important adaptations, or whether we believe, with Geoffrey Miller, that musicality confers an independent reproductive advantage on the genes that produce it, the fact is that such theories have little or no bearing on the nature and meaning of music.[3] The case is exactly like mathematics. It could be that mathematical competence is a by-product of other and more useful adaptations; or it could be that it is an adaptation in its own right. But neither theory tells us what mathematics is, what numbers are, what mathematical truth

[2] Jerry A. Fodor, *Representations: Philosophical Essays on the Foundations of Cognitive Science*, Cambridge Mass., MIT Press, 1981.

[3] Steven Pinker, *How the Mind Works,* New York, Norton, 1997, p. 534. Geoffrey Miller, 'Evolution of Human Music through Sexual Selection', in Nils L. Wallin, Björn Merker and Steven Brown, eds., *The Origins of Music*, Cambridge Mass., MIT Press, 2000.

https://doi.org/10.1017/S1358246114000277 Published online by Cambridge University Press

is, or what mathematics really means. All the philosophical questions remain when the evolutionary account is called in. And the same is true of most of the problems that concern philosophers of music.

Matters are otherwise with the computational theory of the brain. There is no doubt that this has cast light on the understanding of language. And it is not implausible to suggest that, if the computational theory goes some way towards explaining language, it might go some way towards explaining music too. For it reminds us that music is not sound, but sound organised 'in the brain of the beholder'. Musical organisation is something that we 'latch on to', as we latch on to language. And once the first steps in musical comprehension have been taken we advance rapidly to the point where each of us can immediately absorb and take pleasure in an indefinite number of new musical experiences. This recalls a fundamental feature of language, and not surprisingly results from linguistics have been transferred and adapted to the analysis of musical structure in the hope of showing just how it is that musical order is generated and perceived, and just what it is that explains the grip that music exerts over its devotees.

We should recognise here that music is not just an art of sound. We might combine sounds in sequence as we combine colours on an abstract canvas, or flowers in a flowerbed. But the result will not be music. It becomes music only if it also makes *musical sense*. Leaving modernist experiments aside, there is an audible distinction between music and mere sequences of sounds, and it is not just a distinction between *types* of sound (e.g. pitched and unpitched, regular and random). Sounds become music as a result of organisation, and this organisation is something that we perceive and whose absence we immediately notice, regardless of whether we take pleasure in the result. This organisation is not just an aesthetic matter – it is not simply a *style*. It is more like a grammar, in being the precondition of our response to the result *as music*. We must therefore acknowledge that music (or at any rate, tonal music of the kind familiar to the Western listener) has something like a syntax – a rule-guided process linking each episode to its neighbours, which we grasp in the act of hearing, and the absence of which leads to a sense of discomfort or incongruity.

Of course there are things called music which do not share this syntax – modernist experiments, African drum music, music employing scales that defy harmonic ordering, and so on. But from mediaeval plainsong to modern jazz we observe a remarkable constancy, in rhythmical, melodic and harmonic organisation, so much so that one extended part of this tradition has been singled out as 'the common practice' whose principles are taught as a matter of

https://doi.org/10.1017/S1358246114000277 Published online by Cambridge University Press

course in classes of music appreciation. This phenomenon demands an explanation.

Leonard B. Meyer, in an influential book (*Emotion and Meaning in Music*, Chicago 1956), argued that we understand music by a kind of probabilistic reasoning, which endows musical events with varying degrees of redundancy. The common practice has emerged from a steady accumulation of conventions and expectations, which enable listeners to predict what follows from what, and which give rise to the distinctive 'wrong note' experience when things go noticeably astray. This suggestion was taken forward by Eugene Narmour, to produce what he called the 'implication-realization model' of musical structure.[4] And more recently David Temperley has applied Bayesian probability theory to standard rhythms and melodies, in order to 'model' the way in which listeners assign meter and tonality to sequences.[5]

Temperley's work raises three questions: what is a 'model'? When is a model 'adequate' to the data? And what might the discovery of an adequate model show, concerning our understanding and appreciation of music? A model that can be rewritten as an algorithm could programme a computer to recognise (or should we say 'recognise'?) metrical order and key. Such a model can be tested against human performance, and if it successfully predicts our preferences and decisions, it offers the beginning of a theory of musical cognition. It suggests an account of what goes on in the brain, when listeners identify the metrical and tonal structure of the piece they are listening to. And that seems to be the aim of Temperley's reflections, especially in his earlier work, in which he develops a computational system for the analysis of music, and uses that system to represent patterns and sequences that are 'preferred' by habituated listeners.[6]

However, others use the term 'model' more loosely, to mean any way of representing the musical surface that displays the perceived connections among its parts, and which suggests a way in which we grasp those connections, whether or not consciously. In this sense the circle of fifths, chord-sequence analysis and the old charts of key relations are all partial 'models' of our musical experience. They enable us to predict, up to a point, how people will respond

[4] *The Analysis and Cognition of Basic Melodic Structures*, Chicago 1990.

[5] *Music and Probability*, MIT Press, 2007.

[6] See David Temperley, *The Cognition of Basic Musical Structures*, Cambridge Mass., MIT Press, 2001, and also Temperley's web-site, which offers access to the Melisma Music Analyzer, a program developed by Temperley and Daviel Sleator.

https://doi.org/10.1017/S1358246114000277 Published online by Cambridge University Press

to changes of key and to accidentals in a melody, and they also suggest musical 'constants' on which a composer can lean when constructing the harmonic framework of a piece. But they do not aim to reduce musical understanding to a computational algorithm, nor do they offer anything like a complete theory of musical cognition, that will explain how we assemble a coherent musical surface from our experience of its parts. Rather, they describe the surface, by identifying the salient features and the perceived relations between them.

Things would look a little different, however, if we could take the idea of a musical 'syntax' literally. Linguistics attempts to model language use and comprehension in ways that lend themselves to computational analysis. If we could extend to the realm of musicology the advances made in psycholinguistics, therefore, we might be nearer to explaining what goes on, when people assemble the notes that they hear into coherent structures. Inconclusive research by the neuroscientists suggests that 'although musical and linguistic syntax have distinct and domain-specific syntactic representations, there is overlap in the neural resources that serve to activate these representations during syntactic processing'.[7] This – 'the shared syntactic integration resource hypothesis' – would be of considerable interest not only to evolutionary psychology but also to musicology, if it could be shown that the syntactic processes involved in the two cases work in a similar way. The neurological research does not show this. But there is a kind of speculative cognitive science that suggests that it might nevertheless be true, and that a 'grammar' of tonal music could be developed which both resembles the grammar of language, and can also be rewritten as a computational algorithm.

One goal of Chomsky's generative grammar has been to explain how speakers can understand indefinitely many new utterances, despite receiving only finite information from their surroundings. Formal languages like the predicate calculus provide a useful clue, showing how infinitely many well-formed formulae can be derived by recursion. If natural languages are organised in the same way, then from a finite number of basic structures, using a finite number of transformation rules, an infinite number of well-formed sentences could be extracted. Understanding a new sentence would not be a mystery, if speakers were able to recuperate from the string of uttered words the rule-governed process that produced it. Likewise the widespread capacity to latch on to new music without any guidance other than that already absorbed through the ear, could be

[7] Aniruddh D. Patel, *Music, Language and the Brain*, Oxford, OUP 2008, p. 297.

https://doi.org/10.1017/S1358246114000277 Published online by Cambridge University Press

explained if musical surfaces were the rule-governed products of a finite number of basic structures, which might be partly innate, and partly acquired during the early years of acculturation.

Certain aspects of music have been modelled in ways that suggest such a generative grammar. If metrical organisation proceeds by division, as in Western musical systems, then surface rhythms can be derived from basic structures by recursion and also understood by recuperating that process. This is made into the basis of a generative grammar of metrical rhythm by Christopher Longuet-Higgins and C.S. Lee[8]. Others have made similar first shots at grammars for pitch organisation.[9]

Such small scale proposals were quickly displaced by the far more ambitious theory presented by Fred Lerdahl and Ray Jackendoff in their ground-breaking book, *A Generative Theory of Tonal Music* (1983). Their argument is bold, ambitious and detailed, and although things have moved on in the thirty years since the book first appeared, it has lost none of its relevance, and continues to be called upon by musicologists, music theorists and philosophers of music, in order to develop or make use of the analogy between linguistic and musical understanding. Lerdahl and Jackendoff recognise at many points, however, that this analogy is stretched, and that Chomskian linguistics cannot be carried over wholesale into the study of tonal music. Syntax, they recognise, does not in music point towards semantics, as it does in language. Moreover, the hierarchical organisation that Lerdahl and Jackendoff propose is an organisation of individual musical objects, such as notes and chords, and not, as in Chomsky, of grammatical categories (verb, noun-phrase, adverb etc.). There are no grammatical categories in music. Moreover, while we can distinguish 'structural' from 'subordinate' events in music, there is much room for argument as to which is which, and there is no one hierarchy that determines the position of any particular event. An event that is structural from the 'time-span' point of view might be metrically subordinate and also a prolongation of some other event in the hierarchy of tension and release. Still, the various hierarchies identified by Lerdahl and Jackendoff capture some of our firmer intuitions about musical importance. The task

[8] 'The Rhythmic Interpretation of Monophonic Music', in Longuet-Higgins, *Mental Processes: Studies in Cognitive Science*, Cambridge MA, MIT Press 1987.

[9] For example D. Deutsch and J. Feroe, 'The Internal Representation of Pitch Sequences in Tonal Music', *Psychological Review* 1981, **88**, 503–522.

https://doi.org/10.1017/S1358246114000277 Published online by Cambridge University Press

is to show that there are transformation rules that derive the structure that we hear from a more deeply embedded structure, and do so in such a way as to explain our overall sense of the connectedness of the musical surface.

To grasp the point of the generative theory of tonal music it is important to distinguish two kinds of hierarchy. A generative hierarchy is one in which structures at the level of perception are generated from structures at the 'higher' level by a series of rule-governed transformations. Perceivers understand the lower level structures by unconsciously recuperating the process that created them, 'tracing back' what they see or hear to its generative source. By contrast a cumulative hierarchy is one in which perceived structures are repeated at different temporal or structural levels, but in which it is not necessary to grasp the higher level in order to understand the lower. For example, in classical architecture, a columniated entrance might be contained within a façade that exactly replicates its proportions and details on a larger scale. Many architectural effects are achieved in that way, by the 'nesting' of one aedicule within another, so that the order radiates outwards from the smallest unit across the façade of the building. This is not an instance of 'generative' grammar in the sense that this term has been used in linguistics, but rather of the amplification and repetition of a separately intelligible design. It is true that the order of such a façade is generated by a rule, namely 'repeat at each higher scale'. But we understand each scalar level in the same way as every other. You recognise the pattern of the entrance; and you recognise the same pattern repeated on a larger scale in the façade. Neither act of recognition is more basic than the other, and neither depends on the other. In *The Aesthetics of Music*[10] I argue that many of the hierarchies discerned in music, notably the rhythmic hierarchies described by Cooper and Meyer[11], are cumulative rather than generative, and therefore not understood by tracing them to some hypothetical 'source'. In the case of rhythm there are generative hierarchies too, as was shown by Christopher Longuet-Higgins, writing at about the same time as Cooper and Meyer. But it seems to me that, in the haste to squeeze music into the framework suggested by linguistics, writers have not always been careful to distinguish the two kinds of hierarchy. Music, in my view, is more like architecture than it is like language, and this means that repetition, amplification, diminution and

[10] Oxford, OUP, 1997, p. 33.
[11] *The Rhythmic Structure of Music*, Chicago 1960.

https://doi.org/10.1017/S1358246114000277 Published online by Cambridge University Press

augmentation have more importance in creating the musical surface than rule-guided transformations of some structural 'source'.

The place of semantics in the generation of surface syntax is disputed among linguists, and Chomsky has not adhered to any consistent view in the matter. As a philosopher, however, influenced by a tradition of thinking that reaches from Aristotle to Frege and Tarski and beyond, I would be surprised to learn that deep structure and semantics have no intrinsic connection. Language, it seems to me, is organised by generative rules not by chance, but because that is the only way in which it can fulfil its primary function, of conveying information. Deep structures must surely be semantically pregnant if the generative syntax is to shape the language as an information-carrying medium – one in which new information can be encoded and received. Without semantically pregnant deep structure language would surely not be able to 'track the truth', nor would it give scope for the intricate question-and-answer of normal dialogue. A syntax that generates surface structures from deep structures is the vehicle of meaning, and that is why it emerged.

Take away the semantic dimension, however, and it is hard to see what cognitive gain there can be from a syntax of that kind. In particular, why should it be an aid to comprehension that the syntactical rules generate surface structures out of concealed deep structures? This question weighs heavily on the generative theory of music, precisely because (as Lerdahl and Jackendoff recognize) music is not about anything in the way that language is about things or in the way that figurative painting is about things. Indeed, musical organisation is at its most clearly perceivable and enjoyable in those works, like the fugues of Bach and the sonata movements of Mozart, which are understood as 'abstract' or 'absolute', carrying no reference to anything beyond themselves. The 'aboutness' of music, for which we reserve words like 'expression' and 'feeling', is a matter of what Frege called tone, rather than reference.

You might say that a hierarchical syntax would facilitate the ability to absorb new pieces. But this ability is as well facilitated by rules that link surface phenomena, in the manner of the old rules of harmony and counterpoint, or by the techniques of local variation and embellishment familiar to jazz improvisers. What exactly would be added by a hierarchical syntax, that is not already there in the perceived order of repetition, variation, diminution, augmentation, transposition and so on? Perhaps it is only in the case of metrical organisation that a generative hierarchy serves a clear musical purpose, since (in Western music at least) music is measured out by division, and

https://doi.org/10.1017/S1358246114000277 Published online by Cambridge University Press

divisions are understood by reference to the larger units from which they derive.[12]

There is a theory, that of Schenker, which offers to show that harmonic and melodic organisation are also hierarchical, and Lerdahl and Jackendoff acknowledge their indebtedness to this theory. According to Schenker tonal music in our classical tradition is (or ought to be) organised in such a way that the musical surface is derived by 'composing out' a basic harmonic and scalar progression. This basic progression provides the background, with postulated 'middle ground' structures forming the bridges that link background to foreground in a rule-governed way. Musical understanding consists in recuperating at the unconscious level the process whereby the background *Ursatz* exfoliates in the musical surface.

Objections to Schenker's idea are now familiar. Not only does it reduce all classical works, or at least all classical masterpieces, to a single basic gesture. It also implies formidable powers of concentration on the listener's part, to hold in suspension the sparse points at which the *Ursatz* can be glimpsed beneath the surface of a complex melodic and harmonic process. Moreover, it leaves entirely mysterious what the *benefit* might be, either in composing or in listening to a piece, the understanding of which involves recuperating these elementary musical sequences that have no significance when heard on their own.

More importantly, the whole attempt to transfer the thinking behind transformational grammar to the world of music is a kind of *ignoratio elenchi*. If music were like language in the relevant respects, then grasp of musical grammar ought to involve an ability to produce new utterances, and not just an ability to understand them when produced by someone else. But there is a striking asymmetry here. All musical people quickly 'latch on' to the art of musical appreciation. Very few are able to compose meaningful or even syntactically acceptable music. It seems that musical understanding is a one-way process, and musical creation a rare gift that involves quite different capacities from those involved in appreciating the result.

Here we discover another difficulty for theories like that of Lerdahl and Jackendoff, which is that they attempt to cast what seems to be a form of aesthetic preference in terms borrowed from a theory of truth-directed cognition. If understanding music involved

[12] Note, however, that there are musical traditions which measure musical elements by addition and not division, notably the Indian traditions studied by Messiaen. See my 'Thoughts on Rhythm', in *Understanding Music*, London, Continuum, 2009.

https://doi.org/10.1017/S1358246114000277 Published online by Cambridge University Press

recuperating information (either about the music or about the world) then a generative syntax would have a function. It would guide us to the semantically organised essence of a piece of music, so that we could understand what it says. But if music says nothing, why should it be organised in such a way? What matters is not semantic value but the agreeableness of the musical surface. Music addresses our preferences, and it appeals to us by presenting a heard order that leads us to say 'yes' to this sequence, and 'no' to that. Not surprisingly, therefore, when Lerdahl and Jackendoff try to provide what they regard as transformation rules for their musical grammar, they come up with 'preference rules', rather than rules of well-formedness.[13] These 'rules' tell us, for example, to 'prefer' to hear a musical sequence in such a way that metrical prominence and time-span prominence coincide. There are over a hundred of these rules, which, on examination, can be seen not to be rules at all, since they do not owe their validity to convention. They are generalisations from the accumulated preferences of musical listeners, which are not guides to hearing but by-products of our musical choices. Many of them encapsulate aesthetic regularities, whose authority is stylistic rather than grammatical, like the norms of poetic usage.

The formal languages studied in logic suggest, to a philosopher at any rate, what might be involved in a generative grammar of a natural language: namely, rules that generate indefinitely many well-formed strings from a finite number of elements, and rules that assign semantic values to sentences on the basis of an assignment of values to their parts. Nobody, I believe, has yet provided such a grammar for a natural language. But everything we know about language suggests that rules distinguishing well-formed from ill-formed sequences are fundamental, and that these rules are not generalisations from preferences but conventions that define what speakers are doing. They are what John Searle calls 'constitutive' rules. Such rules have a place in tonal music: for example the rule that designated pitches come from a set of twelve octave-equivalent semitones. But they do not seem to be linked to a generative grammar of the kind postulated by Lerdahl and Jackendoff. They simply lay down the constraints within which a sequence of sounds will be heard as music, and outside which it will be heard as non-musical sound. Moreover these constitutive rules are few and far between, and far less important, when it

[13] Likewise, the theory of musical cognition advanced by David Temperley in his earlier work, *The Cognition of Basic Musical Structures, op. cit.,* is formulated in terms of 'preference rules'.

https://doi.org/10.1017/S1358246114000277 Published online by Cambridge University Press

comes to saying how music works, than the résumés of practice that have been studied in courses of harmony and counterpoint.

This brings me to the crux. There is no doubt that music is something that we can understand and fail to understand. But the purpose of listening is not to decipher messages, or to trace the sounds we hear to some generative structure, still less to recuperate the information that is encoded in them. The purpose is for the listener to follow the musical journey, as rhythm, melody and harmony unfold according to their own inner logic, so as to make audible patterns linking part to part. We understand music as an object of aesthetic interest, and this is quite unlike the understanding that we direct towards the day-to-day utterances of a language, even if it sometimes looks as though we 'group' the elements in musical space in a way that resembles our grouping of words in a sentence.

This does not mean that there is no aspect to musical grammar that would deserve the sobriquet 'deep'. On the contrary, we recognise long-term tonal relations, relations of dependence between episodes, ways in which one part spells out and realises what has been foretold in another. These aspects of music are important: they are the foundation of our deepest musical experiences and an endless source of curiosity and delight. But they concern structures and relations that are created in the surface, not hidden in the depths. The musical order is not generated *from* these long-term relations as Schenker would have us believe, but points *towards* them, in the way that architectural patterns point towards the form in which they culminate. We come to understand the larger structure as a result of understanding the small-scale movement from which it derives.

One of the strengths of *A Generative Theory of Tonal Music* is that it emphasizes these long-term relations, and the way in which the listener – especially the listener to the masterworks of our listening culture – hears the music as *going somewhere*, fulfilling at a later stage expectations subliminally aroused at an earlier one. The mistake, it seems to me, comes from thinking that these perceived relations define a hidden or more basic structure, from which the rest of the musical surface is derived. The perceived relations should rather be seen as we see the relation between spires on a Gothic castle. The pattern made by the spires emerges *from* the supporting structures, but does not generate them.

So where does this leave the cognitive science of music? Thanks to Turing and 'information technology' a particular image of mental processes has taken hold in philosophy. According to this image all mental processes of which understanding is a crucial component are syntactical operations, in which the 'logic gates' of the brain

https://doi.org/10.1017/S1358246114000277 Published online by Cambridge University Press

open and close in obedience to algorithms that link input to output in ways that fulfil the cognitive needs of the organism. This powerful image feeds from our own attempts, in computer technology, to 'reverse engineer' the human cortex. A telling instance is provided by the digital image, transmitted from camera lens to computer screen. The image on the screen is composed of coloured pixels, which are themselves generated digitally, from information supplied by the camera, and transferred algorithmically to the screen. There is no miracle involved in this process, nor would it be a miracle if something similar occurred, by way of transferring information from the retina of the eye to the optical centres of the cortex.

However, the image on the screen is not just an array of pixels: it is an image *of* something – of a woman, say. It has the crucial 'aboutness' that has so often eluded the theories of psychology, and which requires an act of what might be called 'semantic descent': the passage from the data to their interpretation. Moreover, although this act of interpretation depends on processes in the brain, it is not the brain but the person who grasps the result of it. The computational theory that explains the transfer of the image from the lens to the screen offers no explanation of what goes on, when that image is interpreted as the image of a woman.

Moreover, the woman in the picture may be imaginary. I may see the woman while believing there is no such woman. And even in the case of photographs, where the assumption is that there is or was just such a scene as the scene portrayed, standing in a causal relation to its image on the screen, the normal person does not think that the things he sees in that image are actually *there* where he sees them, on the screen.

Imaginary objects, like mathematical objects, and other objects that slip without trace through causal networks, pose well-known problems for naturalistic theories of the mind. But they also remind us that we cannot dispense with philosophy. Such objects raise a question about intentionality that must be solved *prior* to cognitive science, if we are to know what shape our cognitive science is to have. We want to know how it is possible for a mental state to be of or about something that is believed not to exist – how a mental state can contain an apprehension of the nothingness of its own object, so to speak. And we want to know how such a mental state can connect with the steady flow of our thoughts and desires, and the developing thread of our emotional life, even though it tells us nothing directly about the world around us. Greg Currie's suggestion, that in the work of the imagination we run our mental states 'off-line', seems to me to be more a description of the problem than

https://doi.org/10.1017/S1358246114000277 Published online by Cambridge University Press

a solution to it.[14] For what is the ground for thinking that the 'on-line' mental state of believing that p is importantly similar to its off-line version of imagining that p? And in what way similar? You don't provide a cognitive science of the mind by using computer science as a source of metaphor.

As I earlier suggested, music is not a representational art form, and musical understanding is in that respect quite unlike our understanding of pictures – it does not involve, as the understanding of pictures involves, the recuperation of an imaginary world. Nevertheless, we do not perceive music simply as sequences of sounds: there is, as Lerdahl and Jackendoff and many others remind us, an act of synthesis, of mental organisation, involved in hearing sounds as music, and this is the equivalent in auditory perception of the moment of 'semantic descent' to which I referred earlier. We do not simply hear the sounds that compose the musical work. We hear sequences of pitched sounds, and we hear *in* those sounds a musical process that is supervenient on the sounds although not reducible to them. I have argued this point at length in *The Aesthetics of Music*. Music involves movement in a one-dimensional space, in which there are fields of force, relations of attraction and repulsion, and complex musical objects like melodies and chords that occupy places of their own. It exhibits opacity and transparency, tension and release, lightness and weight and so on. I have argued that there is an entrenched metaphor of space and movement underlying all these features of music. Some – notably Malcolm Budd[15] – dispute my claim that this metaphor is as deeply entrenched as I suppose it to be. But, as I have argued elsewhere,[16] the alternative view – that the parameters of musical movement and musical space can be described in the literal language of temporal progression – must also concede that musical understanding involves grasping activity, movement, attraction and repulsion, and a host of other phenomena that are not reducible to sequential ordering or to any other physical features of the sounds in which they are heard. Those features are part of what we perceive, when we hear music, and someone who merely hears sequences of pitched sounds – *however accurately* – does not hear music. (You could have absolute pitch and still not hear music. Some birds are like this.)

The great question, then, is what cognitive science can tell us, about hearing and understanding music. I earlier gave arguments

[14] *Arts and Minds*, Oxford, Clarendon Press 2004, chapters 9 and 10.

[15] Malcolm Budd, 'Musical Movement and Aesthetic Metaphors', *British Journal of Aesthetics*, 2003.

[16] *Understanding Music*, London, Continuum, 2009, ch. 4 'Movement'.

https://doi.org/10.1017/S1358246114000277 Published online by Cambridge University Press

for dismissing the view that there is a generative syntax of music. But it is clear that music is organised in the ear of the beholder, and that all those features to which I have just referred, whether or not based in entrenched metaphors, are features of the organisation that we impose upon (or elicit in) the sequences that we hear. So how should the cognitive science proceed? One thing is clear: it cannot proceed simply by adapting cognitive science models from other areas, such as the cognitive science of language. We have to start from scratch. But there is very little scratch to start from, at least in the work of those cognitive scientists who have attended to this problem. Thus Aniruddh Patel, who has made a consistent effort to summarise the relevant findings of neuroscience, begins his discussion of melody from the following definition: melody is 'a tone sequence in which the individual tones are processed in terms of multiple structural relationships'.[17] But what is a tone? Is it identical to a pitched sound, or something that is heard *in* a pitched sound? What kind of 'relationships' are we talking about, and why describe them as 'structural'? You can see in this very definition a host of short cuts to the conclusion that music is processed in something like the way language is processed, and that 'processed' is just the word we need – the very word that suggests the algorithms of computer science. But maybe it is not like that at all. How would we know? Surely it is exactly here that philosophy is needed.

The first thing that a philosopher ought to say is that we understand music in something like the way we understand other art forms – namely, imaginatively. *Pace* Budd and others, I believe that when we hear music we hear processes, movements, and relations in a certain kind of space. This space is what is represented in our standard musical notation, and it is one reason why that notation has caught on: it gives us a clear picture of what we hear, unlike, say, the graph notation used by lutenists or the fret-board notation for the guitar, which give us a picture of the fingers, rather than the tones. If we adhere to the strict sense of 'model' that I referred to earlier, according to which a model is the first step towards a computational algorithm, then it is clear that no model can make use of the phenomenal space that is described by ordinary musical notation. A space in which position, movement, orientation and weight are all metaphors is not a space that can feature in a computer program, or indeed in any kind of theory that seeks to explain our experience, rather than to describe its subjective character. It is a space that is *read into*, *imposed upon*, *elicited in*, sounds when perceived by a

[17] *Music, Language and the Brain*, p. 325.

https://doi.org/10.1017/S1358246114000277 Published online by Cambridge University Press

certain kind of perceiver – one who is able to detach his perceptions from his beliefs, and to put normal cognitive processes on hold (or 'off line', to use Currie's metaphor).

Here is a point at which we might wish to step in with a long suppressed protest against the ambitions of cognitive science. Musicology, we might say, is or ought to be a humanity and not a science. It is not a prelude to a theory of musical cognition, whatever that may be. It is devoted to describing, evaluating and amplifying the *given* character of musical experience, rather than to showing how musical preferences might be tracked by a computer. Hence the one-dimensional pitch space in which *we,* self-conscious and aesthetically motivated listeners, situate melodic and harmonic movement, is the real object of musical study – the thing that needs to be understood in order to *understand music*. But this one-dimensional pitch space is not a space in which the physical events that we hear as music actually occur. An account of 'auditory representations' which offers to explain what goes on when we hear music will therefore not be an account of anything that occurs in that imaginary space. No account of auditory sequences and their 'processing' in the brain will be an account of what occurs in the imaginary space of music.

I shall conclude with a few observations that suggest, to me at least, that the philosophy is needed before the cognitive science can begin, and that the premature desire for an explanation may actually distort our account of the thing that needs to be explained. First, there is an asymmetry in music between the listener's, the composer's and the performer's competence. You can acquire a full understanding of music even though you cannot compose, and even though you cannot perform. I have already touched on this point, but it would surely have to be developed philosophically prior to any attempt at a cognitive science of musical appreciation.

Secondly, there is a kind of freedom in musical perception which parallels the freedom in the perception of aspects in the visual arts, but which is absent from ordinary cognitive processes. In the Müller-Lyer illusion the apparent inequality of the lines remains even after the subject knows that the lines are of the same length – proof, for some psychologists, of the modular nature of our sensory and intellectual processes, which deliver independent information about one and the same state of affairs. In the case of aspect perception, by contrast, appearances can change under the influence of thought, and will change if the right thinking is brought to bear on them. The nude in Titian's *Venus of Urbino* changes appearance if you imagine her to be looking at a lover, a husband, or merely a curious observer. How you see her depends upon how you think of

https://doi.org/10.1017/S1358246114000277 Published online by Cambridge University Press

her size, which in turn depends on how you think of the size of the bed on which she is lying. And so on. In the case of music the structural relations to which Patel refers are multiply adaptable to the needs of musical thought. Melodies change according to our conception of where upbeats end, where phrases begin, which notes are 'intruders' and which part of the flow, and so on. And it is one reason why performers are judged so intently, namely that how they *play* can influence how we *hear*.

Thirdly, we do not hear music as we hear other sounds in our environment. Music is heard as *addressed* to us. We move with it, regard it as calling on our attention, making demands on us, responding to our response. Enfolded within the music there lies an imagined first-person perspective, and to listen with full attention is to relate to the music as we relate to each other, I to Thou. Musical movement is a kind of action, and the 'why?' with which we interrogate it is the 'why?' of reason and not the 'why?' of cause. Hence the imagined space of music is a 'space of reasons', to use Wilfrid Sellars's well known idiom, and what we hear in it we hear under the aspect of freedom. This feature is integral to the meaning of music, and is one reason why we wish to speak of understanding and misunderstanding what we hear, regardless of whether we can attach some separately identifiable meaning to it. No doubt cognitive science will one day tell us much about the forms of inter-personal understanding. But it will have to advance well beyond the theory of auditory perception if it is to complete the task.

Those features, it seems to me, demand philosophical exegesis. They ask us to look at the phenomenon itself, to identify just what makes an experience of sound into an experience of *music*. Only when we have clarified that question, can we go on to ask questions about the neural pathways involved and the way the sounds are 'processed' by them.

But there is, I think, a more important topic that opens here, and one that must be the subject of another article. Even if we came up with a theory about the processing of music, it would not, in itself, be an account of musical understanding. Indeed, it would tell us as little about the meaning and value of music as a cognitive model of mathematical understanding would tell us about the nature of mathematical truth. All the real problems, concerning what music means, why we enjoy it, and why it is important to us, would remain untouched by such a theory. For they are problems about the experience itself, how that experience is profiled in our own first-person awareness, and what it means. Meaning is opaque to digital processing, which passes the mystery from synapse to synapse as a relay

https://doi.org/10.1017/S1358246114000277 Published online by Cambridge University Press

team passes the baton, or as the algorithm passes the image in my earlier example. The crucial moment of 'semantic descent' certainly occurs. But it involves the whole cognitive and emotional apparatus, and achieves an act of understanding of a kind that has yet to find its place in the computational theory of the mind. But here we are in deep water, and there are as many philosophers who will disagree with that last sentence (Fodor, for instance) as there are who will agree with it (Searle, for instance).

Ethics and Public Policy Center, Washington D.C. and
the University of Oxford
rogerscruton@mac.com

https://doi.org/10.1017/S1358246114000277 Published online by Cambridge University Press