

On Evolutionary Explanations of Musical Expressiveness

Matteo Ravasio*

*I am grateful to Stephen Davies, Justine Kingsbury, Justin Horn, Daniel Wilson and Marinus Ferreira for comments and lively discussions on previous versions of this paper.

On Evolutionary Explanations of Musical Expressiveness

Abstract

In this paper, I will examine an evolutionary hypothesis about musical expressiveness first proposed by Peter Kivy. I will first present the hypothesis and explain why I take it to be different from ordinary evolutionary explanations of musical expressiveness. I will then argue that Kivy's hypothesis is of crucial importance for most available resemblance-based accounts of musical expressiveness. For this reason, it is particularly important to assess its plausibility. After having reviewed the existing literature on the topic, I will list five challenges the hypothesis is supposed to meet. Although my list of challenges does not aim at exhaustiveness, I believe that the hypothesis must meet all of the challenges I suggest if it is to work as a cornerstone for a theory of musical expressiveness.

Keywords

Music
Expressiveness
Evolution
Resemblance
Animation



1. Introduction

Recent research on musical expressiveness has proposed a variety of evolutionary explanations to account for our propensity to hear music as expressive of emotions. In this paper I will focus on a particular explanation first proposed by Peter Kivy.¹ I will contend that this evolutionary hypothesis should be distinguished from scientific evolutionary accounts of the mechanisms responsible for the perception of emotional expression in music. A hypothesis similar to Kivy's is accepted, in one form or another, by various contemporary philosophers of music. However, some authors, starting with Kivy himself, have expressed skepticism as to its plausibility. I believe that these doubts are well founded. Given the importance of the proposal in the literature, it is necessary to examine more closely the problems it poses. I will outline five challenges, all of which seem difficult for Kivy's hypothesis to meet, at least at the present stage of elaboration. Although my list does not aim to be exhaustive, I believe that failing to meet these challenges is likely to represent a fatal flaw for the argument in question.

2. Kivy's evolutionary hypothesis

I shall first briefly introduce Peter Kivy's contour theory of musical expressiveness, for it is in response to the problems encountered by such a view that Kivy resorts to his evolutionary story. According to the contour theory of musical expressiveness, music is expressive in virtue of its resemblance to emotional prosody and other expressive behavior— such as the adoption of a certain gait, carriage, or countenance to express an emotional state. This general suggestion is not new; it is indeed as old as Plato, although Kivy prefers to identify the music theorist Johann Mattheson as a closer forerunner.²

Kivy's novelty lies in his intuition that the emotional quality of the music, i.e., its expressiveness, is logically distinct from any actual emotion in the listener or in the composer. That is, although music might arouse an emotion in the listener and/or have resulted from an emotional state the composer was in when he wrote it, the expressive character of the music is independent from both the aroused emotion and the emotion felt by the composer. In this way Kivy distinguishes the contour theory from both the arousal theory of musical expressiveness and the so-called expression theory, which respectively identify the music's expressive character with

On Evolutionary Explanations of Musical Expressiveness

the aroused emotion and the emotion expressed by the composer. Kivy's famous example to illustrate this is the Saint Bernard dog's sad expression: we readily perceive the sad look of the animal in virtue of its similarity to a sad human face; and yet this expressive character is independent of the actual emotional state of the dog as well as of anyone's intention to express one's emotion.³

At this point, Kivy has to face two kinds of worries. *Firstly*, one might argue that the experience of the resemblance in question is not ubiquitous and certainly is not always a conscious one. Although we might point out ways in which music is like people expressing emotions, we are not required to notice them when perceiving the emotions in the music. We do not normally experience music as resembling human behavior, and Kivy is surely not interested in producing a prescriptive account about how we should listen to expressive music.⁴ *Secondly*, if we decide to ground the expressiveness of music in the music's resemblance to expressive utterances and expressive behavior, we must confront the objection according to which music does not resemble cries, moans, carriages and gaits any more than it resembles many other things, such as waves' motion, the fury of the elements during a storm, or "the rise and fall of the stock market or the spirit of capitalism."⁵ But music is clearly not expressive of all these things—although it might in some cases be considered to represent some of them. An analysis resting only on resemblance might consequently miss the target, and clearly such an analysis is the one offered by Kivy so far.

The evolutionary story has the crucial task of defending the contour theory from these objections by (1) providing reasons to believe that there need not be any conscious perception of the resemblance between music and expressive behavior and (2) explaining why, among the many things music resembles, emotional expression is the one we hear in it.

Kivy's suggestion is that we are hard-wired to animate inanimate objects because of evolutionary reasons. As we are likely to mistake for a snake the stick we stumble upon while walking in a wood (we see the snake in the stick in virtue of their similarity), we have an unconscious tendency to animate music and perceive emotional expressiveness in it (we hear the expressive behavior in the music in virtue of their similarity). Because of the primacy of the sense of sight, Kivy argues, the perceptual error in the case of the stick/snake takes the form of a conscious experience. It is important to note how ambiguous Kivy is in the examples he offers for the visual case. In *The Corded Shell* he describes the experience of seeing a human figure

in a spoon or a face in a simplified drawing of a circle with horizontal traits constituting the eyes, nose, and mouth.⁶ I will assume for the sake of the argument that there is no difference between these two cases, although this might be contested. However, in his *Introduction to a Philosophy of Music* he offers the already mentioned example of the stick/snake, which strikes me as clearly representing a different sort of case from the other two situations.⁷ This choice of examples is hardly irrelevant, as I hope to show later.

The sense of hearing, Kivy continues, is comparatively less important for us as a species, and this explains why the perception of emotions in music does not need to be conscious. Whereas we see the stick as a snake and are startled, the perception of the resemblance between music and the emotions does not produce an instinctive response. It might be that things were different for our ancestors: maybe they perceived threatening or friendly utterances in sounds as much as we see friendly or hostile gazes in the natural environment around us; but the course of evolution and the primacy of sight have turned sound animation into a sort of “vestigial relic.”⁸

As we will see more clearly later, Kivy himself came to doubt the plausibility of his contour theory. The conjectural nature of the evolutionary argument is one of the reasons for this change of mind.⁹

3. Ordinary scientific evolutionary hypotheses and the specificity of Kivy’s story

I do not question that expressiveness is related to some evolved mechanism. We have good reasons to believe that the perception of expressive qualities in inanimate objects is grounded in our evolved nature—as most universally shared psychological mechanisms and tendencies arguably are. Let us briefly consider two evolutionary hypotheses of this kind in order to illustrate this last point.

Norman D. Cook has proposed a suggestive ethological explanation for the expressive character of major and minor chords.¹⁰ It is widely known that, at least in Western tonal harmony, major chords are described as cheerful, whereas minor chords are perceived as subdued or gloomy. Cook believes that, if we look at how major and minor chords are approached from situations of musical tension, we might be able to discover the source

On Evolutionary Explanations of Musical Expressiveness

of their expressive character. He observes, drawing on previous work by Leonard Meyer, that any triad made of equally spaced notes sounds tense.¹¹ If we accept this narrow notion of musical tension as triads composed of equidistant intervals, we can see how the resolution of the tension by means of a semitone increase or decrease in any of the notes is going to invariably produce a minor chord whenever the semitone increases, a major chord whenever it decreases. If we look at ethology, Cook continues, we see that humans share with other animals a sort of auditory code according to which increases in the fundamental frequency of a sound are linked with calls of defeat and submission—whereas decreases in frequency are typical of calls communicating victory, dominance, and strength. This frequency code carries over into human language—think of the rising tone of polite requests and of the descending tone of assertions and commands—and of course into music, where we see it operating in the expressive character of major and minor chords. I believe Cook's theory to be implausible for a number of reasons, although it is not my concern here to address the difficulties it faces.

Another evolutionary hypothesis about musical expressiveness is defended by Jaak Panksepp.¹² He proposes an evolutionary explanation for the sad/bittersweet quality of music containing features such as high-pitched crescendos or a solo instrument emerging from an orchestral background. These musical features are often associated with the sensation of chills or musical *frissons*. Panksepp suggests that the reason for this is the similarity between the separation calls typical of various mammals and the musical features under consideration. Listening to such music could therefore activate ancient emotional circuits deputed to the regulation of emotions of paramount importance for social life.

Neither Cook nor Panksepp argues for or against any particular phenomenology of music listening. It might be that, as James O. Young seems to believe, Panksepp's theory could be used to support a resemblance-based theory of musical expressiveness such as Kivy's—although I am skeptical about it.¹³ My point here is merely that there is a way of theorizing the evolutionary mechanism grounding musical expressiveness without committing to phenomenological assumptions about music listening. It is true that both Cook and Panksepp accept some basic first-person characterizations of music—for instance, that we perceive the major mode as happy and the minor as sad. However, the main goal of these theories is to describe from a third-person perspective the connection between facts related to our evolved nature and facts related to our perception of

emotion in the music or to the music's power of emotional arousal. They leave unanswered many of the questions related to the phenomenology of expressive music.

The kind of argument proposed by Kivy is of a more specific sort. The main difference resides in the relation proposed by Kivy between evolved tendencies and the phenomenology of music listening. For Cook and Panksepp, phenomenology is a mere starting point for scientific explanations which do not of themselves amount to phenomenological claims. In contrast, Kivy's claim about the phenomenology of music listening is quite specific: our way of listening to music involves a reference to human expressive behavior, that is, to the sort of behavior that typically accompanies emotions in human beings. More precisely, Kivy believes that we perceive music as expressive because we perceive it—at least subliminally—as an instance of human emotional expression, bodily or vocal.

Kivy's argument would appeal to theorists who believe that the perceptual material provided by the music constitutes the input of a mechanism, the output of which is a phenomenal experience necessarily involving human expressive behavior. However, his hypothesis is not a mere placeholder for future empirical discoveries about the evolutionary links between musical expressiveness and the expression of emotions. It constitutes a specific claim as to how we should interpret those empirical findings: it is in other words a philosophical argument that can be questioned without doubting the relevant underlying scientific facts.

4. Who needs the evolutionary hypothesis?

Needless to say, Kivy's contour theory needs the evolutionary hypothesis if it intends to answer to the two problems presented earlier. As already anticipated, Kivy became skeptical about the viability of his contour theory, also because of the lack of support for the evolutionary hypothesis on which the theory is grounded. He observes: "What evidence, if any, is there for the claim that listeners subliminally hear the analogy, if indeed it exists, between the contour of music and human expression? And even if they do hear it, does that adequately explain our experience of hearing emotions in the music as perceptual qualities? As well, does the phenomenon of seeing things in ambiguous figures—seeing the stick as a snake, or the faces and figures in clouds—transfer to sounds and what we hear (if anything) in

On Evolutionary Explanations of Musical Expressiveness

them?”¹⁴ To these doubts, some of which will be developed and amplified in my own challenges to the evolutionary hypothesis, Kivy adds more general methodological concerns about ‘just so stories’ lacking experimental support.

The interest in outlining challenges for the evolutionary hypothesis comes from the fact that various other philosophers need something akin to Kivy’s concept of animation in order for their theory of musical expressiveness to work. As a first approximation, we could say that all resemblance-based theories of musical expressiveness require a story about animation such as Kivy’s. We shall now examine this issue in more detail.

Jerrold Levinson observes that there is no way to answer the question as to how we come to hear emotion in music, rather than something else, except by appealing to “our disposition to aurally construe the music as an instance of personal expression, perceiving the human appearances in the musical ones, in effect animating the sounds in a certain manner, to use a phrase given currency by Peter Kivy.”¹⁵ A tendency to animate music is also essential to Levinson’s theory of musical expressiveness—the so-called persona theory, according to which hearing expression in music is hearing it as an instance of expression by a musical persona. Moreover, we should notice that Levinson needs a hard-wired disposition such as the one described by Kivy for the same reasons Kivy needs it. On the one hand, Levinson believes that the resemblance between music and human emotional expression has a role in our perception of musical expressiveness. However, as everything resembles everything else, he has to justify why we hear emotions in the music rather than something else. On the other hand, Levinson is inclined to stress that the listener does not need to explicitly imagine a fictional persona expressing herself musically, as this might well only occur “in a back-grounded manner.”¹⁶ An evolved tendency to animate our perceptions could account for the subliminal character of this process, just as it did in Kivy’s case.

Stephen Davies offers a somewhat more cautious and subtle analysis of Kivy’s speculations.¹⁷ He is inclined to think that Kivy demands from his evolutionary story more than is needed. In particular, Kivy struggles to justify an asymmetry between the visual and auditory domain. This asymmetry is not only implausible, but also unnecessary to his argument. It is implausible because perceptual errors such as the stick/snake case happen in perfectly analogous fashion in the auditory case

too—one might be inclined to mistake a click in the dark for the cocking of a gun. It is unnecessary because animation, Davies argues, occurs in both visual and auditory cases in circumstances where there is no possible perceptual mistake involved. When we look at a portrait, we *animate* the portrait and say, for instance, that the portrayed person is looking to her left although we are perfectly aware of having in front of us nothing more than pigments on canvas.¹⁸ Notice that Davies’s revision of Kivy’s argument is more detrimental to its original function than it might at first seem. From an argument that served, among other things, to distinguish expression from resemblance-based representation, we arrived here at a defense of the concept of animation that construes it exactly as a case of pictorial depiction, that is, a case of resemblance-based representation. Animation in the sense Kivy intends is clearly not just the *representation of things that are animate* but rather a *genuine experience of facing a perceptual content as we face something that is animate*.

I will finally note how Davies, who believes that music is expressive in virtue of its presentation of emotion-characteristics-in-appearance, seems to need something like Kivy’s evolutionary hypothesis for the usual two reasons: 1) accounting for the fact that we perceive the music’s expressive qualities without being aware of any resemblance between the music and human expressive behavior and 2) explaining the salient character of the resemblance between the music and human expressive comportment.

James O. Young, in a recent defense of an anti-formalist philosophy of music, has argued that Kivy’s resemblance theory is obviously right and supported by both common sense and empirical research.¹⁹ Although I think Young’s theory is further away from Kivy’s than Young realizes, it is worth noting that his account, at least to the extent that it actually is akin to Kivy’s, needs the evolutionary hypothesis just as the original contour theory did.

5. Criticism of the evolutionary hypothesis

We have already seen how Kivy himself became skeptical about the plausibility of his own evolutionary hypothesis. I shall now briefly summarize the criticism found in the existing literature before going on to list the challenges the hypothesis should meet.

On Evolutionary Explanations of Musical Expressiveness

Anthony Newcomb has expressed two worries: firstly, a notion such as animation does not leave any space for the composer's intention, as it describes expression as a projection of the listener; secondly, Kivy's view cannot account for all the *inanimate* properties we ascribe to the music, such as "references we commonly hear in it to water, glass, fire"²⁰ In *Sound Sentiment*, the second, expanded edition of *The Corded Shell*, Kivy has convincingly dealt with these two objections.²¹ As to the first worry, he rightly observes that the composer's intention is preserved, as the composer intentionally uses musical material that is going to provoke some particular expressive animation. In this sense, the link between the perception of expressive content and the composer's choice is no more problematic than a painter's choice of a cold palette to express a desolated, downcast mood. In answer to the second problem, Kivy observes that his story about how we animate music was meant to explain how we come to hear expression in the music; it is therefore pointless to remark that it cannot account for descriptions of music in terms of fire, glass, and other inanimate substances, for these are clearly not things that music or anything else could possibly express.

Geoffrey Madell has argued that the evolutionary hypothesis is not compatible with Kivy's claim that some musical elements are expressive in virtue of conventional association rather than because of their contour similarity with human expression—a case in point being the expressive character of major and minor chords.²² Madell rightly observes that Kivy requires the notion of animation to distinguish expression from mere resemblance-based representation: music resembles many things, but our evolved tendency to animate makes its resemblance to human expression stand out, triggering our experience of expression in the music. The problem, Madell observes, is that whereas animation requires some sort of resemblance, conventional association does not require any. From Kivy's standpoint, then, it is hard to explain how we perceive the conventional expressive character of, say, major and minor harmonies just as we experience the expressiveness resulting from the animation process. According to Kivy's analysis, expressive conventions are strictly speaking no cases of musical expressiveness at all. Although I find Madell's objection quite compelling, I should note that it is dependent upon the acceptance of a theory of musical expressiveness that exploits both resemblance-based expression and expression based on conventional association. This latter aspect could be rejected and the former could be broadened to explain the cases that conventional association was meant to cover.

Derek Matravers does not offer any direct criticism of the evolutionary hypothesis itself. He notes, however, how Kivy's concept of animation forces him to downplay the role of resemblance in his account of musical expressiveness. If music resembles many things other than those it expresses, then the experience of expressive music is not the awareness of a resemblance but merely our own expressive animation of it.²³ I am unconvinced by this line of reasoning, as one could reply that the evolutionary hypothesis is nothing but a causal story to explain why we are prone to notice certain similarities rather than others. Animation does not need therefore to be severed from the awareness of a resemblance.

6. Five challenges for the evolutionary argument

Rather than offering a defense or a refutation of Kivy's evolutionary hypothesis, I will outline some challenges it should be able to meet in order to be plausible. I will be mostly concerned with Kivy's characterization of the argument, as he offers the most elaborate one. However, in light of the fact that Kivy no longer favors resemblance theories of musical expressiveness, it will be particularly important to assess the weight of the challenges for accounts such as Davies's and Levinson's.

The phenomenological challenge

We have already noticed how Kivy specifies that the perception of the animated content is subliminal in hearing: we do not hear expressiveness in music as we mistake the stick for the snake because sight has a primary adaptive importance for us—a fact that keeps the tendency to construe resemblances between visible things at the conscious level. It seems to me that denying any similarity or convergence between seeing and hearing is a way for Kivy to cover up the fact that he is actually dealing with two different kinds of experience, both of which involve perception but in very different ways.

Things are further complicated by the fact that the examples Kivy offers for the visual case do not fall into the same category as his auditory examples. In fact, I believe that the cases from which, according to Kivy, the evolved animation tendency stems are of a radically different kind from the ones in which we see something in something else (which is what he

On Evolutionary Explanations of Musical Expressiveness

takes musical animation to consist in). If animating the inanimate means seeing a snake where there is only a stick, I agree that there is a reason to think this ability might be an advantage. However, I do not believe this is what happens in the animation of music. Consider, on the one hand, the experience of seeing a snake where there is only a stick and, on the other hand, the experience of seeing a smiling face in the front of a car, a face in moss, or hearing sad human emotional behavior in music. There seems to be more than a mere difference in degree between the first kind of experience and the second. In the first case we are dealing with perception. Because perception is influenced by personal and cultural values, we might perceive one thing rather than another more readily—even to the point of committing perceptual mistakes. I am not seeing the snake in the object I am looking at, as when I see a figure in a moss pattern. I am simply believing that there is a snake, although the perceptual data informing that belief could be differently interpreted and lead me to the right conclusion that there is no snake.

We can further stress the difference between perceptual error and aesthetic animation if we consider when and how we attempt to justify our experiences of things. I can make sure the stick is not a snake by hitting it with my walking stick, throwing a stone at it, or stomping on the ground to make the snake/stick go away. But in seeing a face in the grille of a car or hearing emotion in music is not a mistake per se—therefore we need not do anything to try to justify our perception. The latter is an actual case of seeing- or hearing-in. Mistaking a stick for a snake is not. If there is something like a rationale for the seeing-in experience, it would consist in pointing out which perceptual aspects of the object (e.g., the grille of a car) are relevant to our experience of “seeing-in” (e.g., an upward curve)—albeit with the understanding that someone else could experience the same object quite differently. To sum up, one can objectively prove or disprove a perceptual error; but what I call “seeing- or hearing-in” is a ‘double aspect’ experience which can be more or less successfully supported by the perceptual qualities of the object under consideration but is otherwise subject to voluntary control.

Kivy is of course aware that we do not make any type of perceptual error when experiencing music as expressive. He is thus aware of the distinction just proposed although he seems to blur it by offering instances of *both kinds of perceptual act* as examples of animation. Once the difference is clouded in such a way, it is easy for him to conclude that the evolutionary tendency underlying perceptual mistakes straightforwardly applies to

seeing-in experiences in which we see something animate in something inanimate.

How is all this relevant to the evolutionary story? The crucial point is that Kivy seems to believe that it is the first kind of experience (stick/snake) which originates the latter (hearing sadness in music). But the two experiences exhibit the phenomenological differences which I have presented above. By conflating the two cases, Kivy masks his need to explain how the evolved tendency to commit certain kinds of perceptual errors rather than others has generated our seeing-in abilities. Accepting my line of reasoning does not amount to denying that the seeing-in experience could have some other evolutionary origin. But Kivy's failure to bridge the gap between perceptual error and seeing-in renders his evolutionary story incomplete.

One might reject my phenomenological challenge from two perspectives: 1) we may accept the phenomenological distinction between the two cases but find that the tendency to animate the inanimate could have developed in one type of situation and then, once established, work for the other situation as well in a by-product fashion; 2) we may believe the distinction is only a matter of degree.

It is hard know how to address the first position. The fact that Kivy offers examples of 'animation' so different from one another as the case of the stick/snake and the one of the spoon seen as a human figure might indicate that this is the strategy he would follow if he were to defend his evolutionary story from my phenomenological challenge. We should note that Kivy later admitted and discussed the ambiguity of his own stick/snake example.²⁴ He recognizes the difference between the case of the stick/snake and the case of music which invokes no startle mechanism. He then offers further reasons to treat the aural stimulus as something that results in a subconscious animation and because of its comparatively lower survival value—the startle mechanism is now in the background, as it were. But this clearly ignores the fact that the startle mechanism is still in place in aural perception (as in Davies's gun example). Therefore I maintain that the actual difference is not between aware and unaware startle mechanisms but between two different types of experience—as suggested by the phenomenological challenge. The abilities involved in perceptual errors (from which the tendency to animate plausibly evolved) are different from those involved in seeing-in and hearing-in; and this should be evident from my phenomenological description of the two experiences.

On Evolutionary Explanations of Musical Expressiveness

The adaptation challenge

If we reject the phenomenological challenge for the second reason and consider the distinction between the two cases only a matter of degree, we can then question the idea that the tendency Kivy described does indeed have an adaptive value. Thus far I have offered no reason to doubt that the tendency to animate the inanimate is valuable from an evolutionary point of view.

If we claimed that the two experiences are different only in degree, such that the animation in the stick/snake case and in the spoon/humanoid case are at opposite ends of a continuum, then we may reasonably hold that the stick and the snake resemble each other more than the front of a car and a smiling face or a wooden spoon and a human figure resemble each other, while still maintaining that the cases are not qualitatively different. However, if the ‘animation tendency’ is reinforced by natural selection to such an extreme degree that it begins to generate ‘byproducts’—like seeing faces in the moss—then it could start to work against its own potential as an evolutionary adaptation. If there is no limit to how weak the resemblance can be between what we animate and the animation we perceive, then this ‘wild animation’ would only lower our chances of survival.

Kivy supposes that our ancestors heard emotional expression in sounds as we see snakes in sticks. But we clearly need to keep the animation under control if it is to be useful. If Kivy decided to defend his thesis by claiming that the animation tendency only works *up to a point*, the thesis would lose its purpose. For if the difference between the stick/snake and the musical case were only to be found in the relative strength of the resemblances between objects and animated perceptions; then because music bears nothing but weak resemblances to other things, the perception of expressiveness in music would be a very weak version of the stick/snake experience (closer to car/smile) and thus contribute little to survival. In sum, if the naturally selected tendency always works, we lose the tendency to animate; if it works only in clear cases like the stick/snake, we lose its power to explain the musical case.

Both the phenomenological challenge and the adaptation challenge are related to the phenomenology of expressive music. Do they represent a serious source of concern for Davies’s appearance emotionalism and for Levinson’s persona theory? An account such as Levinson’s might be immune from such phenomenological worries, as it considers the music’s

resemblance to expressive behaviors as one of the *grounds* of musical expressiveness rather than as part of its phenomenology. Davies, as we have seen, believes that Kivy's worries about the animation process are not justified. We animate things all the time, just as when we see a person painted on a flat surface. The problem with this comparison is that it eludes the worries about musical animation by pointing to experiences that are considerably different from musical ones: when we look at a picture, we are typically well aware of the depicted object, which normally is the main object of attention. Kivy's original goal was to provide an evolutionary explanation for the characteristically backgrounded, subliminal character of the experience of expressive behavior in music.

The sense modality challenge

Sound animation is relevant to evolutionary adaptation only insofar as animated sounds retain informational meanings in the life of an individual. It is clear how the tendency to perceive a growl in what is not a growl might be relevant for the survival of the individual. Recall also Davies's example of a click in the dark perceived as the cocking of a gun. To draw a connection with the visual stick/snake case: better to run away from a stick than to grab a snake thinking it is a stick.

However, in many musical cases the perceived resemblance is cross-modal: on Kivy's account, music can resemble salient *bodily movements* of people who express certain emotions.²⁵ How could such a case of synesthetic animation have adaptive value? In order to answer this question, proponents of the evolutionary hypothesis need to provide a more elaborate explanation than Kivy's. His hypothesis might be able to justify our propensity to perceive the similarity between music and vocal emotional expression; but it does not seem to be capable of explaining why we perceive resemblances between music and visual emotional correlates such as bodily movements. A sound can of course suggest a movement in the sense that the movement can be related to it as a physical cause, but this is of no help to Kivy's idea. The problem posed by this challenge could be particularly pressing for accounts such as Davies's, which stresses the analogy between music and bodily behavior rather than vocal behavior.

On Evolutionary Explanations of Musical Expressiveness

The obliquity challenge

It is commonly admitted—even by those who do not accept the cognitivist resemblance-based account of musical expressiveness—that descriptions of musical expressiveness which emphasize the resemblance between music and expressive gestures are legitimate. However, too close a resemblance to expressive behavior is normally considered deleterious to the purpose of musical expressiveness. The music may well be shaken by rhythmical variations and brisk tempo changes or leap through the tonal space. But when instrumentalists start to imitate notated gestures too closely or a singer's voice is broken by sadness or screams of pain, it is no longer musical expressiveness that we are dealing with but rather some sort of musical representation or theatrical device. The perception of musical expressiveness seems then to be inhibited by an extreme articulation of the resemblance between music and the expressive gesture. The music's analogy with expressive gesture is thus a subtle one, and needs, as it were, to be kept concealed. I call this the *obliquity condition* of musical expressiveness.

I will passingly note that Schopenhauer seems to have been the first philosopher to notice this. His account of musical expressiveness is deeply embedded in his metaphysics of the Will: it is because of music's relation with the Will that music acquires its capacity to embody feelings. The mystery of musical expressiveness is represented by the unintelligible possibility of music to represent what is by definition beyond any possibility of representation, namely the noumenal essence of the world, the Will. The fundamental source of musical meaning is therefore thoroughly different from the phenomenal world, hence Schopenhauer's skepticism about the use of 'painterly' tricks in music. He writes: "But the analogy discovered by the composer between these two [the music and the stirrings of the will] must have come from the immediate knowledge of the inner nature of the world unknown to his faculty of reason; it cannot be an imitation brought about with conscious intention by means of concepts, otherwise the music does not express the inner nature of the will itself, but merely imitates its phenomenon inadequately."²⁶

This interpretation of musical expressiveness is bound to cast a further doubt on the role of animation. Resemblance is no foe to animation; it is actually its greatest ally. The animation of the stick, which becomes to our eyes a dangerous snake, is ultimately more vivid and convincing than the animation of the spoon/human, and this is because a stick can have a

considerable number of visual features in common with a snake. However, the obliquity condition implies that resemblance in the musical case has to be kept within precise limits if we are to avoid the risk of transforming music into the kind of expressive pantomime that seems to be at odds with musical expressiveness. A central feature of animation, namely the role resemblance plays in it, does not seem to fit well with the role resemblance has in musical expressiveness. This constitutes a further reason to doubt that animation plays a role in musical expressiveness, at least in the sense outlined by Kivy. Davies's and Levinson's accounts, to the extent to which they do not qualify the sort of resemblance that is supposed to ground the experience of expressive music, are also liable to these challenges.

The style challenge

An evolutionary hypothesis such as Kivy's stresses the role of the automatic animation of the musical contour. The hard-wired mechanism to which Kivy resorts would seem to predict a strong agreement among listeners from different musical cultures as to the expressive character of a given piece of music. However, Kivy warns us that things might be more complicated than that. He takes as an example the Indian tradition of *rāgas*. Each *rāga* is associated with a specific *rāsa*, that is, an emotional state that the *rāga* is supposed to evoke. Kivy has doubts about the expressive transparency of this style. He goes as far as to say that "To the uninitiated ear, every *rāga* presents about the same mood: a kind of exotic stupor."²⁷ Recent psychological results seem to show that he was wrong: untrained Western listeners are rather good at identifying the emotion a *rāga* is supposed to express and arouse.²⁸ Because of his pessimistic view about the cross-cultural transparency of expressive music, Kivy needs to explain how the hard-wired response could fail to produce common responses in human beings sharing the same evolved traits. His solution to this issue is ingenious: animating a musical contour might well be an automatic response, but the perception of the musical contour is not, as it requires experience with the musical system in question. Kivy suggests an elegant analogy with the visual domain: "We cannot expect the Western ear to hear, *ab initio*, the expressive contour in Indian music, any more than we can expect an Australian aborigine to see expression in Rembrandt's sketch of a face—not because he fails to read the expressive conventions, but because, to begin with, he fails to read the pictorial ones."²⁹ The problem with this solution is that it increases even more the distance between the hard-wired

On Evolutionary Explanations of Musical Expressiveness

case of the stick/snake and the musical case. Not only is the animation of the inanimate something different than the startle mechanism at work in the case of perceptual mistakes: it is also something that needs training to become operational.

I am not aware of any discussion devoted by Levinson to the issue of the cross-cultural transparency of musical expressiveness. Davies, on the other hand, has devoted considerable attention to the issue, suggesting that the expression of emotions is cross-culturally consistent to an extent that encourages a mild optimism as to its transparency, at least while we wait for further empirical research.³⁰

7. Conclusion

In this paper I have described two main kinds of evolutionary arguments that can be offered in attempt to explain musical expressiveness. The first kind, which has not been my concern here, is represented by scientific hypotheses about the relationship between music as a perceptual object and the widespread human tendency to describe it as expressive as well as, in some cases, to be moved by it. The hypothesis considered in this paper is of a different sort, as it entails some constraints as to the phenomenology of music listening. Peter Kivy, who originally proposed it, became himself skeptical about its plausibility. However, it is important to assess its value, as the hypothesis surfaces in various ways in recent literature on musical expressiveness. I hope to have pointed to some challenges that the argument must meet if it is to work as a cornerstone for any theory of musical expressiveness. My list does not aim at exhaustiveness. I take it, however, that an evolutionary hypothesis such as the one considered here would have serious chances of being successful if it could meet all of the challenges I have presented.

Notes

- 1 Peter Kivy, *The Corded Shell: Reflections on Musical Expression* (Princeton: Princeton University Press, 1980).
- 2 Peter Kivy, "Mattheson as Philosopher of Art," *The Musical Quarterly* 70, (1984): 248-265.
- 3 Kivy, *The Corded Shell*, 12.
- 4 *Ibid.*, 57.
- 5 *Ibid.*, 62.
- 6 *Ibid.*, 58.
- 7 Peter Kivy, *Introduction to a Philosophy of Music* (Oxford: Oxford University Press, 2002).
- 8 *Ibid.*, 43.
- 9 *Ibid.*, 46-47.
- 10 Norman D. Cook, *Tone of Voice and Mind: The Connections between Intonation, Emotion, Cognition and Consciousness* (Amsterdam: John Benjamins Publishing, 2002).
- 11 Leonard B. Meyer, *Emotion and Meaning in Music* (Chicago: University of Chicago Press, 1956).
- 12 Jaak Panksepp, "The Emotional Sources of 'Chills' Induced by Music," *Music Perception: An Interdisciplinary Journal* 13 (1995): 171-207.
- 13 James O. Young, *Critique of Pure Music* (Oxford: Oxford University Press, 2014), 17.
- 14 Kivy, *Introduction to a Philosophy of Music*, 46.
- 15 Jerrold Levinson, "Musical Expressiveness as Hearability-as-expression." In *Contemporary Debates in Aesthetics and the Philosophy of Art*, ed. Matthew Kieran (Oxford: Blackwell, 2006), 192-204, at 197.
- 16 *Ibid.*, 193.

On Evolutionary Explanations of Musical Expressiveness

- 17 Stephen Davies, *Musical Meaning and Expression* (Ithaca: Cornell University Press, 1994), 258.
- 18 *Ibid.*, 260.
- 19 Young, *Critique of Pure Music*, 11.
- 20 Anthony Newcomb, "Sound and Feeling," *Critical Inquiry* 10 (1984): 614-643, at 618-619.
- 21 Peter Kivy, *Sound Sentiment: An Essay on the Musical Emotions* (Philadelphia: Temple University Press, 1989), 177-180.
- 22 Geoffrey Madell, *Philosophy, Music and Emotion* (Edinburgh, Edinburgh University Press, 2002), 25.
- 23 Derek Matravers, *Art and Emotion* (Oxford: Oxford University Press, 1998), 122.
- 24 Kivy, *Sound Sentiment*, 172.
- 25 Kivy, *The Corded Shell*, 52-54.
- 26 Arthur Schopenhauer, *The World as Will and Representation*, trans. E. F. J. Payne (New York: Dover, 1969), 263. A similar idea was previously expressed by Charles Avison in his *Essay on Musical Expression*, first published in 1752. Avison writes: "And, as Dissonance and shocking Sounds cannot be called Musical Expression; so neither do I think, can mere Imitation of several other Things be entitled to this Name, which, however, among the Generality of Mankind, hath often obtained it. Thus the gradual rising or falling of the Notes in a long Succession, is often used to denote Ascent or Descent, broken Intervals, to denote an interrupted Motion, a Number of quick Divisions, to describe Swiftiness or Flying, Sounds resembling Laughter, to describe Laughter; with a Number of other contrivances of a parallel Kind, which it is here needless to mention. Now all these I should chuse to stile Imitation, rather than Expression; because, it seems to me, that their Tendency is rather to fix the Hearers Attention on the Similitude between the Sounds and the Things which they describe, and thereby to excite a reflex Act of the Understanding, than to affect the Heart and raise the Passions of the Soul." (Charles Avison, Pierre Dubois and William Hayes, *Charles*

Avison's Essay on Musical Expression. With Related Writings by William Hayes and Charles Avison [Farnham: Ashgate, 2004], 24). Notice how in this passage there seems to be, *in nuce*, a criticism of a cognitivist perspective on musical expressiveness, such as resemblance theories. The perception of a resemblance is only an “act of the Understanding,” and expression is taken by Avison to require something more—or something different. A few pages later, Addison makes a similar point: “The power of Music is, in this respect, parallel to the power of Eloquence: if it works at all, it must work in a secret and unsuspected Manner” (Ibid., 28).

27 Kivy, *The Corded Shell*, 89.

28 Laura-Lee Balkwill and William Forde Thompson, “A Cross-cultural Investigation of the Perception of Emotion in Music: Psychophysical and cultural Cues,” *Music Perception: An Interdisciplinary Journal* 17 (1999): 43-64.

29 Ibid., 91.

30 Stephen Davies, “Cross-cultural Musical Expressiveness: Theory and the Empirical Program,” in *Musical Understandings and other essays on the philosophy of music* (Oxford: Oxford University Press, 2011), 34-46.

On Evolutionary Explanations of Musical Expressiveness

References

- Avison, Charles, Dubois, Pierre and Hayes, William. *Charles Avison's Essay on Musical Expression. With Related Writings by William Hayes and Charles Avison*. Farnham: Ashgate, 2004.
- Balkwill, Laura-Lee and Thompson, William F. "A Cross-cultural Investigation of the Perception of Emotion in Music: Psychophysical and cultural Cues." *Music Perception: An Interdisciplinary Journal* 17 (1999): 43-64.
- Cook, Norman D. *Tone of Voice and Mind: The Connections between Intonation, Emotion, Cognition and Consciousness*. Amsterdam: John Benjamins Publishing, 2002.
- Davies, Stephen. *Musical Meaning and Expression*. Ithaca: Cornell University Press, 1994.
- . "Cross-cultural Musical Expressiveness: Theory and the Empirical Program." In *Musical Understandings and other essays on the philosophy of music*. Oxford: Oxford University Press, 2011.
- Kivy, Peter. *The Corded Shell: Reflections on Musical Expression*. Princeton: Princeton University Press, 1980.
- . "Mattheson as Philosopher of Art," *The Musical Quarterly* 70 (1984): 248-265.
- . *Sound Sentiment: An Essay on the Musical Emotions*. Philadelphia: Temple University Press, 1989.
- . *Introduction to a Philosophy of Music*. Oxford: Oxford University Press, 2002.
- Levinson, Jerrold. "Musical Expressiveness as Hearability-as-expression." In *Contemporary Debates in Aesthetics and the Philosophy of Art*. Edited by Matthew Kieran (Oxford: Blackwell, 2006), 192-204.
- Madell, Geoffrey. *Philosophy, Music and Emotion*. Edinburgh: Edinburgh University Press, 2002.
- Matravers, Derek. *Art and Emotion*. Oxford: Oxford University Press, 1998.
- Meyer, Leonard B. *Emotion and Meaning in Music*. Chicago: University of Chicago Press, 1956.

Newcomb, Anthony. "Sound and Feeling," *Critical Inquiry* 10, (1984): 614-643.

Panksepp, Jaak. "The Emotional Sources of 'Chills' Induced by Music." *Music Perception: An Interdisciplinary Journal* 13 (1995): 171-207.

Schopenhauer, Arthur. *The World as Will and Representation*. Translated by E. F. J. Payne. New York: Dover, 1969.

Young, James O. *Critique of Pure Music*. Oxford: Oxford University Press, 2014.

On Evolutionary Explanations of Musical Expressiveness