Beyond Wittgenstein’s Musical Formalism

Marina Bakalova
ISSK, BAS
marina.bakalova@gmail.com

Abstract: In his scattered remarks on music, Wittgenstein adopts a rule-based account of musical understanding also known as musical formalism. This is the view that understanding a piece of music does not presuppose matching the musical sequence with one’s own mental states. It entails instead, recognizing the proper use of conventions and theoretical terminology of a given musical system, such as the Western musical system. I argue that Wittgenstein’s musical formalism is misguided. It raises important questions that it has no resources to answer. The opposite view, according to which musical understanding is tightly related to cognitive, emotional or imaginative states of the listener, can answer these questions better. We have to take into account that if Wittgenstein were to witness the cognitive revolution he would most probably have changed his mind concerning scientific potential of psychology and our grasp of the mental realm. In particular, he may have to accept that certain psychological phenomena account for the appropriateness of our inter-relational comparisons.

Key words: Wittgenstein, musical formalism, cognitive revolution

In his later period, Wittgenstein famously suggests that communicative activities (as well as thinking, understanding
etc.) need not be considered as correlated to mental acts or states. Instead, on his view, the ability to follow the practical conventions of communication is, by itself, necessary and, together with our well-informed contextual associations, sufficient to generate the meaning of our communicative acts. In his scattered remarks on music, Wittgenstein adopts the same position in relation to musical expression, and musical understanding. In *PI* (pp. 181–83), for instance, he claims that sensation is irrelevant to the meaning of a musical phrase. Hence, understanding a piece of music does not presuppose matching the musical sequence with one’s own mental states. It entails instead, recognizing the proper use of conventions and theoretical terminology of a given musical system, such as the Western musical system. I will call this position “rule-based account of musical understanding, or “musical formalism” together with Ahonen (2005).

In this article, I will argue that Wittgenstein’s musical formalism is misguided. It raises important questions that it has no resources to answer. The opposite view, according to which musical understanding is tightly related to cognitive, emotional or imaginative states of the listener, can answer these questions better. I offer a positive mentalist alternative to musical formalism claiming that the primary function of musical expression is to articulate phenomenal character of real or imaginary mental states thereby making them no less than directly perceptible to the listener. I claim, in contrast to what Wittgenstein thought in his later work, that relatively stable meaning is encoded by some of the aesthetic properties of musical works. This is demonstrated by an experimentally proven cognitive phenomenon, the phenomenon of inter-modal congruence that can be seen as central to our capacity of understanding and appreciating art (Green 2007, Ch. 7). Such a mentalist approach to musical understanding is compatible with the idea that knowledge of musical theory and musical culture
generally increases our understanding of musical meaning. More importantly, in certain cases, inter modal congruence accounts for the appropriateness of our inter-relational comparisons which formalists think singlehandedly determine the meaning of musical works.

1. The statement F

A central message of Wittgenstein’s PI is that there is nothing like stable meaning. Despite Wittgenstein’s opposition to any sort of theory in this period, Garry Hagberg (2017) meaningfully systematizes his anti-essentialist ideas about language and their counterpart in aesthetics. He coins Wittgenstein’s late position Relational Determination of Content. According to this view, each meaningful expression has relationally embedded properties that the interlocutor is supposed to get if she is to understand the expression. Hagberg shows in intricate detail how this view underlies Wittgenstein’s remarks of music. Applying Hagberg’s interpretation to Wittgenstein’s anti-mentalism yields the conviction that it is impossible to imagine our own states without using public language and public conventions. Instead we can suggest comparisons between publically accessible objects, but the content of our mental states always remains private and inaccessible to others. Thus the combination between relational determination of content and Wittgenstein’s anti-mentalism entails the following statement:

F: The constitutive relational properties of a meaningful expression or artwork have no context-invariant meaning determined by one’s mental states.

In what follows we will take F as our target statement. First, we will try to show in more detail how F underlies Wittgenstein’s view of musical understanding and then we will raise our objections against it.
2. **Music, Facial Recognition, and Menschenkenntnis**

Music typically expresses mental states, such as emotions or feelings. Wittgenstein seems to agree with that:

‘The aim of music: to communicate feeling.’

But then he continues:

"Connected with this: We may say correctly “his face has the same expression now as previously” even though measurement yielded different results on the two occasions."

The analogy between music and facial recognition is considered central for Wittgenstein’s reflections on music and it gives specificity to the statement F when applied to music. There is a single core, around which the interpretation of this analogy circulates. To start with, the tricky element is that the second part of the quote stands in contrast to the first part. The idea is that facial expressions are not reflections of mental states. Instead, facial expressions function as a part of many public criteria of applying certain mental terms. The same holds for music by analogy. Yes, it “expresses” feelings, according to Wittgenstein, but in the musical gesture, so to speak, there is nothing beyond certain ways of applying mental terms. Expression then is something like a craft. It is not a medium in which the artisan “pours” one’s own feelings. It is a medium in which the artisan demonstrates her craftsmanship or ability to work with public signs to suggest as-if-feelings, *similes*, to provoke associations of the audience. The resulting work can be sophisticated or not based on that ability. Although musical communication is nothing like making you understand my feelings, music can nevertheless transmit understanding of humankind (Menschenkenntnis).

In 1948, Wittgenstein explicitly wrote: “Understanding music is a manifestation of human life” (*CV*, 80). This idea goes beyond the notorious rule following, or at least beyond a single language game. Béla Szabados nicely puts it: for Wittgenstein...
“music is not alone, but it reverberates and resonates with the whole field of our language games—with our artistic and social practices.” (Szabados 2006, p. 657). Eran Guter stresses that Menschenkenntnis is not a body of theoretical knowledge like psychology. Rather, it is a skill, or a set of skills, which some people have a more intuitive grasp of than others (Guter 2017, 236). Guter thereby links Wittgenstein’s understanding of music to the Romantic idea of depth that is typically German. The surprising element related to that interpretation is that Wittgenstein purportedly sees the kinship of music to the inexhaustible domain of interrelated comparisons as concealment. Guter explains that in the context of Wittgenstein’s final writings musical expression is grounded in imponderable evidence; evidence that cannot be recognized or explained by reference to rules, yet it is accepted by those who are experienced with the infinite variation of human physiognomy:

“Understanding music is always in mediis rebus, enmeshed with human life, and that is why we begin by misunderstanding, that is, by noticing that there is something concealed, or somehow hidden or held back” (Guter, 2017, 243)

That concealment, according to Guter leads to revelation. There are two problems with this idea. First it would be quite hard, if ever possible to explain how concealment leads to revelation. Admittedly, there is something about that in the facial recognition analogy. Despite your painful screaming I cannot perceive your pain, or understand how your pain feels to you. It is concealed for me. However, your screaming in pain makes me suffer with you; your scared face can makes me worry. Such cases can sometimes yield to revelation, but not because of concealment in my opinion, neither because one makes educational guesses based on imponderable evidence. Rather, it is because of emotional contagion. So secondly, facial gestures do not usually work by concealment. They are systematic indications of particular
emotions. They are not meant to conceal but to reveal.

On Wittgensteinean picture, understanding life through music is actually enmeshed with how things feel to us but this aspect is contingent to an extent of being irrelevant to what we are supposed to understand, the humankind in its complexity. I disagree with that irrelevance claim. I certainly agree that the depth of human life lies in concealment most of the time, but I am prone to think that the moments when we listen to music are kind of exceptional. I do not believe that music suffers from cognitive and expressive impotence to share with sufficient clarity the complexity or depth of human life. I hold the straightforward belief that music reveals something from that depth period. Very often music reveals the phenomenal content of our mental lives; it articulates such content in detail and I will show how in the last part of this article. Misunderstanding of musical expression (which accompanies concealment) is not a standard situation when we listen to music. Usually it is clear what kind of emotion, mood or feeling music expresses.

Roger Scruton (2017) offers a more promising interpretation of Wittgenstein’s attempt to explain music by analogy to facial recognition. He suggests that Wittgenstein does not give a full-fledged argument about musical understanding, but he nevertheless makes a contribution by connecting music to facial expressions. Scruton acknowledges that Wittgenstein’s conception of understanding facial expressions seems to stay at the level of recognizing expressions on another’s face. However, Scruton wants to take this notion further by claiming that one can gain access to the other person’s first-person knowledge, i.e. knowledge of ‘what it is like’. He thinks that we can adopt the other person’s perspective by imagining her expressed state beyond the immediate musical Gestalt’. (pp. 7–9). I think, and I will try to show that the relation of adopting the other person’s perspective can be even stronger than imagining. It can be a
Scruton’s interpretation of Wittgenstein was criticized by Ahonen (2005). She points out that for Wittgenstein the connection between meaning and understanding is not one of making a connection between the sign and its meaning (an object, event, or property) ‘in the understanding of those who use the sign’. Rather, the meaning of the sign is its rule-governed use in its context, and understanding it is nothing but the ability to follow these rules. (p. 514).

From what we have seen so far, I don’t think that Wittgenstein’s analogy between music and facial recognition is really productive in explaining how we understand music. Understanding music is not like reading faces. It is more than that. It enables us to hear and actually perceive certain emotions in their phenomenal detail. Unlike facial recognition, music makes the phenomenal content of one’s state available to the listener. Precisely in this sense, it can be a revelation.

3. Internal and External Interrelational Comparisons

Let us try a more general strategy to explain musical understanding based on Wittgenstein’s ideas. Gary Hagberg suggests that relational interconnections are constitutive not only of the meaning of our sentences, but also of musical themes (and artworks more broadly). This is the way in which art and language are analogous without being the same thing.

In order to better explain the constitutive nature of relational interconnections, Hagberg distinguishes between internal and external relational interconnections. Internal interconnections pertain to the structure of the artwork. To illustrate that, he gives an example of a poem by Frida Schanz, quoted by Wittgenstein himself:

Foggy day. Gray autumn haunts us.
Laughter seems tainted;
the world is as silent today as though it had died last night. In the red-gold hedge fog monsters are brewing; the day lies asleep. The day will not awaken. (Hagberg 2017, 66)

Wittgenstein notes that he does not know if the first words ‘Foggy day’ are the title or the first line of the poem. If “gray” is the first word, Hagberg claims, then the poem becomes trivial and it “changes the rhythm of the whole poem.” In the two cases we have the same words, but different meaning depending on the relation of the first sentence to the rest of the poem. This example shows that internal relations are constitutive of the artwork and indispensable. Hagberg points out that because of the indispensable nature of its internal structure, the meaning of an artwork cannot be transmitted in any other way but by reiterating the very same artwork. Any attempt to translate it would be a violation of these internal, constitutive interrelations and would entail a substantial loss of its meaning.

There is a problem lurking here, that is only indirectly linked to my argument. The problem springs from the possibility of perfectly recreating the internal relations of an artwork in an indistinguishable forgery that would prompt the same relational interconnections. In this case, a question follows: is the artwork really indispensable? This question is difficult to answer if we stay at the level of interconnected internal structure only. My answer would be that there is a mental part expressed in the original that was plagiarized in the forgery. But, according to Wittgenstein’s treatment this mental part escapes the audience’s gaze in both cases. So, internal relations may fail to account for the indispensability of an artwork after all.

Hagberg points out that Wittgenstein discussed also external interconnections which contribute to aesthetic appreciation and interpretation of artworks. They include comparing artistic styles, observing relations and kindships between various
cultural phenomena, etc. Hagberg claims that not all relational comparisons are legitimate. He excludes “a category of relational interconnections that go too far, or develop without restraint or without an awareness of and respect for prior established relations.” (68) Hagberg’s suggestion is intuitively clear, but it seems to me that there are two cases in which we can legitiately draw relational comparisons without restrain, and without respect for previously established connections. One can 1) read some artworks in an appropriately creative way whatever that means and 2) one can have an insight about an artwork that nobody has had before. Such comparisons would be illegitimate on Hagberg’s proposal. We need a clearer criterion that would, at the same time, unbound the countless possibilities for a legitimate artistic expression and legitimate understanding thereof.

4. A mentalist alternative to Wittgenstein’s musical formalism.

In the last chapter of his book “Self-Expression”, Mitchell Green (2007) develops a theory of artistic expression. Green points out that the main role of artistic expression is to show how a feeling feels. This is possible, he claims, thanks to the phenomenon of inter-modal congruence (also known as cross-modal congruence) (Green 2007: 178-182). According to this phenomenon, some sensations within one sensory modality seem to bear more of an affinity to some sensations within another sensory modality than to others. Examples of the phenomenon include: people systematically relating high pitch with bright light and low pitch with gloomy light. The same holds between sensations, moods, emotions, and all states with phenomenal character. For example, the major chord sounds systematically cheerful to us, whereas the minor chord sounds sad. Other examples that Green cites concern people thinking that yellow is more like the sound of a piccolo than it is like the sound of an oboe; that the smell of sulfur is more
like rough than it is like smooth; that the taste of lemon is more like the minor chord C–E flat–G than it is like the major chord C–E–G, etc. (179) The congruence, Green notices, holds in an irreversible way:

“(S)ome kinds of inter- as well as intrapersonal inversions do not seem possible. It is difficult, for instance, to see how there could be an interpersonal inversion as between pain and pleasure. This would require that the experience that I feel upon cutting my hand with a knife is like the experience you feel upon stroking velvet. Likewise, we can rule out the possibility that a minor chord sounds sad to me but happy to you.” (184)

Green provides a methodologically insightful explanation of why and how the congruence occurs. He offers a multi-dimensional explanatory model for understanding the congruence. Green suggests that our sensations, emotions and moods may be described along a number of dimensions, including the following three: intense/mild, pleasant/unpleasant, dynamic/static. Green limits his talk to these three basic dimensions for simplicity, having in mind that the theory allows for introduction of unlimited number of such dimensions. Hence, he introduces the idea of a three-dimensional space. He asks us to imagine the above mentioned dimensions as placed in a coordinate system creating a three-space. (179) He further assumes that we perceive phenomenal characteristics as regions in the three-space. There are few possibilities to locate states in that space. First, imagine a joy which stands phenomenally in a particular intersection between the three dimensions (say rather intense, rather pleasant and relatively dynamic). It will be represented with one point in the three-space, an intersection between the three parameters. We will call these points “regions”. Alternatively, a state might have phenomenal characteristics only along the intense-mild dimension, but not in the other two dimensions. That state will
be mapped as a single point on one dimension only. Finally, a state might be two-dimensional, for instance pleasant and mild but neither dynamic nor static. For instance, the taste of a smooth cream. Then, it will be mapped with two dots on two regions along the intense-mild and pleasant-unpleasant dimensions.

The idea of the three-space provides a powerful explanation of how a stimulus can express states with phenomenal character. Compare our example of a joy which stands phenomenally in a particular intersection between the three dimensions (intense, pleasant and dynamic) with an auditory experience that occupies the same or very close particular intersection of these three dimensions (i.e. an intense, pleasant and dynamic sequence of sounds). The two experiences feel the same because they are congruent in the three-space. Because of their congruence, that sequence of sounds can express that kind of joy, a listener can hear it as that kind of joy. To put it more formally, let us distinguish between an expressive stimulus “ES” and an expressed state, “E”. We can plausibly assume that when an ES overlaps with a region or a set of regions in the three-space that E occupies, it expresses E. In the case of precise mapping between the regional spaces of E and ES, ES will be quite informative about the qualitative character of E. If E is too complex, presumably ES can give us an idea of how E feels even if it does not map all the regions of E. Certainly there is an issue of how much overlapping is enough to make a qualitative character of a state plausibly recognizable.

Let us turn back to music. Music is an art that develops through time. So, if we want to use the three-space theory to explain musical expressiveness we will have to assume that musical expression draws a trajectory of expressive stimuli or regional activations in a multidimensional space and in a multidimensional time. Musical trajectory (MT) is the musical counterpart of what I just called ES with its progression in time. Imagine a state E, developing in time from point A to point D,
and activating different regions in the three-space at different moments in time: ABC and D. MT expresses E by activating the same three dimensional regions as E. Thus MT enables a listener to experience through her perceptual imagination the phenomenal character of a mental episode developing in time.

I would like to provide some examples of how Western musical theory allows composers to “describe” mental attitudes by using the three-space theory. I will give just a few examples based on tempo, dynamics and harmony.

Musical tempo allows one to place a musical fragment along static-dynamic dimension. The standard expressive tempo forms vary form grave (very, very slow) to largo (very slow) to adagio (quite slow) to andante (a walking pace) to moderato (moderate) to allegro (fast and cheerful) to vivace (lively) to presto (very fast) and finally to prestissimo (fastest). The overall shape of the melody going upward, downward, or remaining static also contributes to that.

Another chapter from the musical theory, dynamics, allows composers to express states along intense-mild dimension ranging from pianissimo to fortissimo. Directions to change dynamics either suddenly or gradually, on the other hand, enable ranging of intensity through time. Here the standard possibilities are crescendo, decrescendo or diminuendo, sforzando, etc.

What about pleasant/unpleasant dimension? Music can express different pleasant states such as joy, liberation, triumph, romance etc. On the other hand, it can express sadness, drama, tragedy, horror, anxiety, etc. and other unpleasant emotions, feelings and moods. There are various ways to suggest an unpleasant emotion in a piece of music. One way is to accelerate tension, typical of emotions like anxiety, distress, drama, and horror. This can be done by increasing the density of the harmony, i.e. by harmonic accumulation. The effect can be achieved also through making the rhythm denser as well as through gradually increasing the
and activating different regions in the three-space at different moments in time: ABC and D. MT expresses E by activating the same three-dimensional regions as E. Thus MT enables a listener to experience through her perceptual imagination the phenomenal character of a mental episode developing in time.

I would like to provide some examples of how Western musical theory allows composers to "describe" mental attitudes by using the three-space theory. I will give just a few examples based on tempo, dynamics, and harmony.

Musical tempo allows one to place a musical fragment along static-dynamic dimension. The standard expressive tempo forms vary from \textit{grave} (very, very slow) to \textit{largo} (very slow) to \textit{adagio} (quite slow) to \textit{andante} (a walking pace) to \textit{moderato} (moderate) to \textit{allegro} (fast and cheerful) to \textit{vivace} (lively) to \textit{presto} (very fast) and finally to \textit{prestissimo} (fastest). The overall shape of the melody going upward, downward, or remaining static also contributes to that.

Another chapter from the musical theory, dynamics, allows composers to express states along intense-mild dimension ranging from \textit{pianissimo} to \textit{fortissimo}. Directions to change dynamics either suddenly or gradually, on the other hand, enable ranging of intensity through time. Here the standard possibilities are \textit{crescendo}, \textit{decrescendo}, or \textit{diminuendo}, \textit{sforzando}, etc.

What about pleasant/unpleasant dimension? Music can express different pleasant states such as joy, liberation, triumph, romance, etc. On the other hand, it can express sadness, drama, tragedy, horror, anxiety, etc., and other unpleasant emotions. There are various ways to suggest an unpleasant emotion in a piece of music. One way is to accelerate tension, typical of emotions like anxiety, distress, drama, and horror. This can be done by increasing the density of the harmony, i.e., by harmonic accumulation. The effect can be achieved also through making the rhythm denser as well as through gradually increasing the density of the facture to \textit{tutti}. We encounter such a way of building tension in “The Death of Tybalt” from the ballet \textit{Romeo and Juliet} by Prokofiev, where the timpani beat and the full capacity of the orchestral \textit{tutti} create the feeling of strong tension and tragedy. Another example is the 4th part of Beethoven’s \textit{Pastoral Symphony} “Thunder Storm”, where we can notice an increase of density of the sound and rhythm, and the use of \textit{crescendo} representing anxiety in the face of a developing storm. Sadness falls in a phenomenally different group of unpleasant states. It is silent, and slow. Sadness can be represented by using minor harmony, timbre in the middle to low register, and articulation in legato.

On the other hand, there are various ways to express pleasant emotions. For example, the choice of a theme is quite important: a well-connected, uninterrupted melody, predominantly in legato, consonant in terms of intervals and harmony creates a feeling of pleasure. For such an effect, melody should have certain diversity that helps the development of the phrase. In other words, it should not be boring. Appropriate articulation and diversity of tempo also contribute to expressing pleasant emotions.

These are just some basic possibilities. A composer can use various expressive means simultaneously to place a musical fragment in a single/multidimensional region, and thus to express and articulate the phenomenal character of a state developing in time. Western musical system alone is a well-developed and structured palette of such expressive means.

Applied to music, Green’s multi-space model has a significant advantage over Wittgenstein’s formalist theory. First of all, it provides adequacy conditions for ascribing expressive content to music, i.e, for musical understanding. In many cases, \( x \) can be expressive of \( p, q, y, \) or \( z \) depending on the associations of the listener. Nevertheless, all these ascriptions can be adequate to \( x \) if they are congruent with \( x \). “Being adequate to \( x \)”, according to the
three-space theory means occupying the same region in the three-space that \( x \) occupies. So, the listener can adequately associate \( x \) with any possible state that occupies the same region in the three-space that \( x \) occupies. Adequacy is the right term here and, of course, it is different from truth. Yet, in some cases, the musical context enables us to form an opinion around one particular state that is the most plausible referent of the expression.

This model does not exclude inter-relational comparisons to play a role in our understanding of musical works. These comparisons, when strictly aesthetical, would also be subject to the above mentioned adequacy conditions. In order to be adequate, they must be based on congruence with the musical work. For instance, Wittgenstein’s statement that Brahms did in full rigor what Mandelson did only half rigorously (Wittgenstein 1931) would be adequate if based on certain congruencies between Brahms and Mandelson.

In conclusion, the development of cognitive science after Wittgenstein shows that his doubt of the science of psychology was premature. At the same time, it is highly probable that he would have changed his mind were he to face the cognitive revolution. So, I think that we should not be afraid to reinterpret his ides in the light of the new discoveries about human mind.

Reference:
The three-space theory means occupying the same region in the three-space that \( x \) occupies. So, the listener can adequately associate \( x \) with any possible state that occupies the same region in the three-space that \( x \) occupies. Adequacy is the right term here and, of course, it is different from truth. Yet, in some cases, the musical context enables us to form an opinion around one particular state that is the most plausible referent of the expression. This model does not exclude inter-relational comparisons to play a role in our understanding of musical works. These comparisons, when strictly aesthetical, would also be subject to the above mentioned adequacy conditions. In order to be adequate, they must be based on congruence with the musical work. For instance, Wittgenstein’s statement that Brahms did in full rigor what Mandelson did only half rigorously (Wittgenstein 1931) would be adequate if based on certain congruencies between Brahms and Mandelson.

In conclusion, the development of cognitive science after Wittgenstein shows that his doubt of the science of psychology was premature. At the same time, it is highly probable that he would have changed his mind were he to face the cognitive revolution. So, I think that we should not be afraid to reinterpret his ideas in the light of the new discoveries about human mind.

Reference


