2022

A Unified Theory of Musical Meaning

Luis Verdin

Claremont McKenna College

A Unified Theory of Musical Meaning

Submitted to
Professor Gabbrielle Johnson

By
Luis Verdin

For
Senior Thesis
Spring 2022
April 25th, 2022

Acknowledgments

I want to thank Claremont McKenna College for all of the resources and opportunities they have provided throughout my college career, and the CMC Department of Philosophy more specifically; the academic rigor and enthusiasm of many of the professors in this department has played a central role in making me the thinker I am today. I'd also like to express gratitude to all of the musicians and philosophers who have most inspired me to take on this project—there are too many of which to name here. A special thank you to my friends, family, and mother in particular, for being there for me when it matters most and supporting my growth. I love you all. Most importantly, I want to thank my thesis advisor, Professor Johnson, for all of the continued guidance, patience, wisdom, and support throughout my thesis-writing process. Working with you has not only been a pleasure and privilege, but also one of the highlights of my academic experience.

Abstract

In this thesis, I present a novel theory of musical meaning. This theory posits a complementary relationship between the theories described in Lerdahl and Jackendoff's *Generative Theory of Tonal Music* and Arnie Cox's *Music and Embodied Cognition.* Each of these theories explain particular aspects of musical meaning (semantic and grammatical, respectively), though I argue that by unifying these theories into a broader framework, they can explain more about musical meaning than they could individually. This unification is performed via the novel theory I present: the Analogical Argument. This argument suggests that Lerdahl and Jackendoff's *Generative Theory of Tonal Music* is theoretically analogous to Noam Chomsky's theory of Generative Linguistic Grammar (Chomsky, 1966). Given the success of Chomsky's theory, as well the cognitive approach he employs more generally, in explaining how we construe linguistic meaning, we should expect similar if not analogous processes to be responsible for the construal of musical meaning. Thus, the Generative Theory of Tonal Music is sufficient for explaining how the musical meanings explained by Arnie Cox's *mimetic hypothesis* are cognized so as to give rise to the emergent musical meaning that is characteristic of musical experiences.

Table Of Contents

We will have achieved much for scientific study of aesthetics when we come, not merely to a logical understanding, but also to the certain and immediate apprehension of the fact that the further development of art is bound up with the duality of the Apollonian and the Dionysian, … their continuing strife and only periodically occurring reconciliation.

—Nietzsche, *The Birth of Tragedy from the Spirit of Music,* 1872

Section 1: Introduction

Music is practiced in all known cultures and throughout history. Underlying this sociological truth is a philosophical puzzle: why, at an individual level, do we find music so enjoyable and, above all, meaningful? We can be moved to tears or feel intense elation, or feel as if we've experienced pure bliss or the transcendence of the self, upon hearing certain compositions. Music has the ability to act as a doorway to alternative states of consciousness, and it can sometimes seem as if it allows us to feel others' feelings. Mysteriously, music has the capacity to be expressive without the use of images, words, and other representations; it is unique among the arts in this regard. This lack of representational content then figures in our phenomenology: we hear music as being expressive of feelings, but face great difficulty in identifying the people and states of affairs that we generally take to be metaphysically prior to feelings, in the music. Though the feelings we have in musical experience are felt to be as real, animated, vivid and meaningful (if not more) as any of the other feelings we have nonetheless. The feelings that can be invoked by music not only seem to vary infinitely, but also have the potential to arise with an overwhelming intensity. It is far from clear how mere sounds have this potential.

While a comprehensive answer to these questions is well beyond our current grasp, developments in science and philosophy have offered insights that grant us the ability to make progress in understanding music. Specifically, Arnie Cox's *mimetic hypothesis* (as presented in *Music and Embodied Cognition*) and Lerdahl and Jackendoff's *Generative Musical Grammar* each provide explanations of two essential components of musical cognition: semantic and grammatical components respectively. In what follows, I explain the meaning we hear in music by unifying these theories into a broader theoretical framework that explains more about musical meaning than each theory could individually.

Section 2: Background

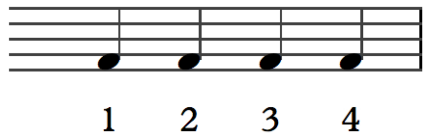
Before tackling the question of musical meaning, it will be necessary to get clear on what we mean by 'music'.

The term 'music' is often ambiguous between the sounds that constitute music, and the affective states we have when hearing these sounds, or both. To avoid such ambiguities in the current project, I'll regard *musical stimuli* as sets of pitch-time events characterized by temporal/rhythmic and tonal/harmonic elements; I'll regard *musical experience,* on the other hand, as what we experience when hearing musical stimuli.

Musical stimuli, or sets of pitch-time events characterized by temporal/rhythmic and tonal/harmonic elements, can be defined in physical terms as sound wave propagations. The two defining features of pitch-time events are of course pitch and time, though both features are already inherent in our definition of sound wave propagations. All sound waves propagate at particular frequencies, and this frequency just *is* pitch; they are also all situated in time. Thus, the term 'pitch-time events' highlights those aspects of sound wave propagations that are most relevant to music, namely pitch and time.

Rhythmic and melodic/harmonic elements of pitch-time events can be realized when we have multiple pitch-time events. Rhythm refers to patterns and regularities in the temporal distribution of pitch-time events. For example, this rhythmic formalization indicates that all four pitch-time events are of the same length, and are therefore each temporally equidistant from their adjacent pitch-time events:

Figure 1

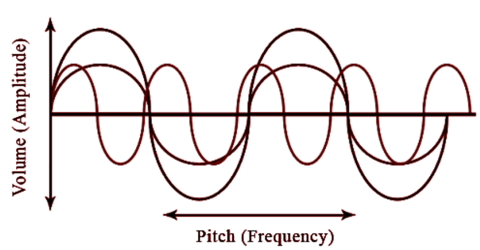


Melody refers to patterns and regularities in sequences of pitch-time events with respect to pitch. Melody is exemplified by the following:

Figure 2

Harmony roughly refers to the extent to which different and simultaneous pitch-time events can 'clash' or 'agree' with each other; we can further ground this in physical terms by reducing this 'clashing' (dissonance) or 'agreement' (consonance) to the extent to which the differing frequencies of different pitches align or interfere with each other:

Figure 3



We can now return to the other concept associated with 'music', namely *musical experience*, or what we experience when hearing musical stimuli. When we listen to musical stimuli, we standardly experience a series of vivid and salient affective states, though sometimes we don't. The extent to which such affective states are invoked in musical experience is seemingly dependent on factors such as how much attention we pay to the music. We will reserve the term 'musical experience' to simply refer to what we experience when we hear musical stimuli, regardless of the salience of the affective states they invoke.

Musical meaning, on the other hand, will refer more specifically to those affective states that are not only invoked in musical experience, but also particularly vivid and salient. Musical meaning refers to the sense we may have in musical experience that what we hear is moving, exciting, sad, angry, *meaningful.*

Notice that all of the properties of musical stimuli thus described are objective and physically defined. The pitch of pitch-time events can be reduced to the wavelength, or distance between the peaks of a sound wave. The temporal component of pitch-time events can be reduced to the time between when the physical source of a sound first resonates so as to produce sound waves, and when it ceases.

Where musical stimuli are objective, musical meaning is subjective. That is, physical features of sounds are publicly observable, while musical meaning is not, as it is experienced subjectively, and our subjective experiences are not publicly available. While we could in principle 'observe' or hear the same musical stimuli as others, we can't 'observe' their mental states or definitively know their character, and it is in this sense that they are not publicly available; it is in this sense that they are subjective.

Though this is not to say that objective, publicly available facts of the world are totally irrelevant to the character of musical meaning. Musical meaning is mentally constructed in such a way that objective features of musical stimuli guide our imaginations, not quite unlike experiences of color: While we subjectively experience color, we know that its qualitative character is not objectively confirmable in the same way objective qualities like pitch or temporal duration are. That said, we can also recognize how the appearance of certain colors in our perceptions is systematic: the colors we perceive are determined by physical properties of what we are perceiving (the frequency of the light waves it reflects), and how the visual stimuli these objective properties give way to (reflected light waves) are processed by our perceptual cognitive systems, which consist of rule-governed computational constraints on how we process visual stimuli.

Thus, we have an explanation for the systematicity of color perception: objective physical features of things in the world determine the stimuli we receive when observing them, and rule-governed computational constraints determine how we arrive at the color (a qualitative thing) we ultimately perceive.

Though in music, we have yet to find a satisfactory explanation for how we go from the objective things (pitch-time events) to the subjective, qualitative things (musical meaning). We want a story to explain this. Relatedly, we should want to explain why certain kinds of chords, compositions, genres, etc. (which are all essentially particular kinds of patterns and regularities in relations among pitch-time events) reliably invoke the same feelings in us. We can refer to this seeming determination of musical meanings by processes that mediate their construal as 'systematicity'. As discussed, the systematicity of color was explained by accounting for our perceptual cognitive systems. Might we expect to find something playing an analogous role for musical meaning?

In what follows, I present a novel theory of musical meaning. The cornerstone of this theory is what I call the 'Analogical Argument'. The core insight of the Analogical Argument is straightforward: Musical meaning is analogous to linguistic meaning, just as standard theories of linguistics following Chomsky regard linguistic meaning as stemming from (1) basic meaningful elements and (2) how these basic elements are related. So too, I argue, can we regard musical meaning as stemming from (1) basic meaningful elements and (2) how those basic elements are related. This prompts two distinct questions:

> **Q1**: How do pitch-time events invoke affective states at all?
>
> **Q2**: In virtue of what do sequences of pitch-time events systematically invoke the affective states (musical meaning) that they do?

Borrowing from Arnie Cox's *mimetic hypothesis* (Cox, 2016) and Lerdahl and Jackendoff's *Generative Theory of Linguistics* (Lerdahl and Jackendoff, 1983), I argue for answers to these questions. First, in response to Q1, I present the mimetic hypothesis in Section 3 which says that particular pitch-time events invoke affective states in virtue of the fact that they resemble sounds we associate with motor actions, which we in turn associate with affective states. Call this mimetic hypothesis 'EMC' (short for embodied musical cognition). By explaining how pitch-time events can invoke affective states, EMC answers Q1; it explains how there can be musical meaning at all. Following this exposition, I will consider possible weaknesses of EMC and evaluate its contributions to the current project in Section 4. Next, in

Section 5, I present Lerdahl and Jackendoff's theory of *Generative Musical Grammar* (GMG). Finally, I argue that Q2 can be answered by the Analogical Argument, the theory that will be presented in Section 6. In Section 7, I elaborate on the implications the Analogical Argument has for considerations of variations in musical meaning among listeners.

This theory posits an interaction between GMG and EMC to explain musical stimuli's systematic invocation of affective states. Further, the Analogical Argument is 'analogical' due to its claim that there is an analogical relationship between GMG, and the processes described in Noam Chomsky's Theory of Generative Linguistic Grammar. This analogy lies in GMG's mediating our relation of pitch-time events so as to give rise to what I will call emergent musical meaning in essentially the same way that the cognitive processes posited in the Theory of Generative Linguistic Grammar mediate our relation of terms so as to give rise to emergent linguistic meaning. We should think GMG can fulfill an analogous role to the processes posited in the Theory of Generative Linguistic Grammar because they both posit a grammar—a set of constraints on how we relate constitutive elements of complex linguistic expressions (GLG) or complex combinations of pitch-time events (GMG). Thus, in explaining how we relate pitch-time events, GMG provides a story of how the musical meanings secured by EMC could be related to each other so as to give rise to emergent musical meaning.

We are now ready to begin our investigation. Though before proceeding, I'd like to make a few preliminary remarks. Firstly, given the central role that affective states will play in the Analogical Argument, I will provide an account of what affective states are, how they interact with other mental states, and answer other relevant questions. Secondly, I'd like to elaborate on why our motivating questions warrant attention, explain how the Analogical Argument is fit to answer these questions, and discuss how the contributions of this project could be situated within the broader music philosophical literature. I begin by providing a metaphysical account of affective states.

Affective states are phenomenal, non-intentional mental states. Here, 'phenomenal' indicates that affective states have a phenomenal character; there is a feeling of what it's like to have particular affective states. This feeling can vary with respect to different phenomenological dimensions. For example, we can experience the affect associated with getting stung by a bee and notice that it is not pleasurable (one phenomenological dimension), while also prominent in our awareness in its onset (another phenomenological dimension) and diminishing in intensity over time (another). The 'non-intentional' component of affective states indicates that affective states do not have representational content; they are not *about* anything in particular. Contrast this with states of belief, which *are* intentional; they are about particular things. For example, I may believe that tigers have black stripes; this belief is about tigers. Though there is no such 'about-ness' to be found for affective states.[1] Given that there is no representational content to affective states, affective states consist exclusively in their phenomenal character.

Further, given that affective states are defined by their phenomenal character, they are conscious states by definition. This is because we can only come to know the character of an affective state through experiencing them, and experience entails conscious awareness. Though to be sure, we can attend to affective states to varying extents. In some cases, we may attend to affective states so minimally that they may be left out of a verbal report of what one is feeling at a given moment, or avoid one's own awareness. Though this should be distinguished from the

[1] One might call this claim into question: the pain of the bee's sting is about the bee, or the sting itself. This seems intuitively plausible, but we should distinguish between the feeling of what it's like to get stung by a bee, and possibly simultaneous mental states such as the perceptual belief that one has been stung by a bee. We may have mental states such as the perceptual belief just described that do have an 'about-ness', but these are distinct from the affective state itself which does not have an 'about-ness'. To be intentional in the relevant sense is to represent some thing or state of affairs in the world. In the affective state associated with getting stung by a bee itself (as well as other affective states), we don't find such representations; affective states could not include representations in principle because they lack the conceptual content that would be required for this.

affective states not being conscious. In order for affective states which we attend to minimally to be included in a verbal report of what one is feeling, or to reach the purview of one's awareness, one would only need to direct their attention towards (attend to) it; it would be a matter of noticing 'what is already there', as it were.

Affective states are of course one kind of mental state we can have, among many. For most (if not all) non-affective mental states, there are accounts we can give of the causal relationships mental states have to other mental states, sensory registrations, and behavioral dispositions. A behavioral disposition is a proclivity towards performing certain behaviors if one is presented with particular sensory stimuli. To give one example of such an account of a causal relationship among mental states, I may have a belief that tigers are dangerous animals. If I also have a desire to avoid dangerous animals, then these states can causally give way to a belief that I should avoid tigers, as well as a behavioral disposition to run if I see a tiger.

I contend that affective states do not have what might be called 'bi-directional' causal relationships with other mental states. That is, affective states can be caused by sensory registrations (e.g. getting stung by a bee), and they can be caused by other mental states (e.g. the invocation of positive affective states as the result of recalling a pleasant memory), but exert no causal influence of their own.

Though this is not to overlook the remarkable extent to which affective states coincide systematically with mental states that *do* have causal influence on other mental states. This coincidence can be so reliable that it may appear *as if* an affective state causes another mental state, though a proper analysis of the mental processes at play in such situations reveals this not to be the case. For example, I may feel the affect associated with a feeling of hunger, and it may seem as if this affect is what causes a desire to visit my fridge. But rather, it is a state of needing food that is responsible for the transition to a state of desire (to visit my fridge). The affective state simply coincides with this transition. I propose that changes in behavioral dispositions only come about through transitions among non-affective mental states. Thus, if affective states have no causal influence over other mental states, then by extension, they can not themselves invoke behavioral dispositions.

Though, as I'll explain in Section 3, *musical affect* is uniquely associated with certain behavioral dispositions by way of EMC, which claims that affective states are invoked by musical stimuli in virtue of associations we have between sounds, affective states, and motor actions. Again, musical affective states, like affective states more generally, can't themselves invoke behavioral dispositions. Though they need not cause behavioral dispositions in order to be associated with them. Not unlike how affective states can coincide systematically with mental states, affective states can coincide systematically with certain behavioral dispositions in musical experience, and it is in this sense that they are associated.

There are two possible concerns with the account of affect I've presented here that are worth addressing. The first concern is about the potential that appealing exclusively to affective states has for explaining all of the feelings we have in musical experience. It may seem as if there is an obvious difference between the affect associated with getting stung by a bee, and what Peter Kivy has called 'garden variety' emotions such as ecstasy, resentment, triumph, jealousy, etc. which seem relatively complex. I contend that such complex feelings are the result of complex relations among basic feelings (Nussbaum, 2007). There are some affective states that are rooted in our senses and can be experienced universally (see Section 7), and are in this sense basic. In the theory of musical meaning I will present, I argue that the invocation of such basic affective states in musical experience can be explained by EMC. Further, GMG can explain how we generate structural descriptions of music. This basically means that GMG explains how we relate musical stimuli, and the basic affective states that come with them, together such that they give way to complex affective states.

Another possible concern about the account of affect I've presented is that it contradicts the widely held intuition that affective states in fact have causal influence over other mental states. More specifically, this account contradicts the intuition that affective states are what motivate our desire to have musical experiences; it seems intuitively and obviously correct to say that one listens to their favorite song repeatedly because they *enjoy* this musical experience. Though not unlike experiences of hunger, I argue that the affective states we encounter in musical experience simply coincide with mental states that have causal influence over other mental states. Presumably, this account will seem especially unsatisfying when applied to music because there doesn't seem to be anything we *get* from musical experiences aside from the experience (and the affective states that come with it) itself. Contrast this with cases such as those where one experiences hunger; while it also seems counterintuitive that the affect of hunger does not cause any transition among mental states, we can still make sense of a causal story about why I ultimately visit my fridge. It is not the mere feeling of hunger that motivates my visiting my fridge, but rather the state of needing food. Satisfying my need for food has obvious practical benefits, and this seems to render a causal story of why I visit my fridge rational.

There is much room for speculation about what an analogous practical benefit might be in a musical context. Some possible answers include music's ability to strengthen social relationships and secure common meaning (Dissanayake, 2000), or how music involves the practice and strengthening of our own capacities for motor simulation (Nussbaum, 2007). There are principled reasons to believe that musical experiences could have a practical benefit(s) that renders their pursuit rational, and indeed, such hypotheses have already found some empirical support. This said, these practical benefits would serve to demonstrate how this account of affect is no more problematic when applied to music than when it is applied to standard examples of affective states such as hunger.

Finally, it's worth mentioning here that I will be limiting my investigation to what is sometimes referred to as 'pure music' in the music philosophical literature. Pure music is music that is not about anything, or has no representational content. This is in contrast to 'program music', which does have representational content; this can be the result of a musical piece including lyrics, having a suggestive title, being written to accompany a play, etc. As has been explained, musical meaning is essentially affective, and affective states are essentially non-intentional. Thus, in the interest of isolating and explaining musical meaning, we will exclude music with representational content, or 'program music', and its possibly confounding influence from our investigation.

With this groundwork set, let us now consider why our motivating questions warrant attention, how the Analogical Argument is fit to answer these questions, and how the contributions of this project could be situated within the broader music philosophy literature.

To further clarify the weight of the puzzle at hand, let us remember that musical experience refers to any instance where we hear musical stimuli and construe musical meaning (which consists in vivid and salient musically invoked affective states), but also those where we don't construe musical meaning. It goes without saying that there are myriad possible musical experiences, though those experiences that give way to musical meaning are most philosophically interesting because the mental states associated with them resist simple explanation. To take one example of a musical experience that does not give way to musical meaning, we can imagine a situation where music is playing quietly, but one is too engrossed in a good conversation to engage with the music. The effect that music would have on its listener in this case is not mysterious: there might be a mere sensory registration and/or basic cognition of the pitch-time events that would not involve the invocation of vivid and salient affective states[2]. We can explain how such states are realized in totally causal terms, and can in fact even build computers that realize them.

In contrast, what *is* mysterious is how we get from physical stimuli to vivid and salient feelings. There exist no such causal stories to explain how these states emerge from, or systematically relate to, musical stimuli. Thus, we should focus our investigation on mental states/situations that display this mysterious invocation as much as possible. This is why explaining musical meaning will be the overarching goal of this project.

Such questions are by no means new. Musical expression has long been the object of philosophical inquiry, and I consider the current project to be an extension of the literature on musical expression[3]. I hope to make novel contributions to this area of investigation by offering a theory of musical expression that can explain and unify two 'worlds' or 'components' of musical experience within a single theoretical framework.

So far as I can tell, the vast majority, if not all, of the theories of musical meaning that have been presented hitherto primarily contribute explanations of only one of the two components of musical experience I have alluded to (the grammatical and the affective/semantic components of musical experience) or fall short of sufficiently explaining how they must

[2] This is not to say that this musical experience couldn't involve any affective invocation at all. Though it is to say that this musical experience would not give way to musical meaning unless the affect that is invoked is vivid and salient, which is not the case.

[3] Though I adopt a slightly different terminology than is traditional in this literature; what philosophers have by-and-large referred to as 'musical expression', I refer to as 'musical meaning' with the intention of foregrounding the role that grammatical processes play in the construal of musical meaning.

interact. In the philosophical domain, many of the theories that have been presented deal only in the affective/semantic component. To consider two prominent examples, persona theories (theories arguing that music is expressive/meaningful in virtue of our imagining a persona whose emotions are being expressed through the music) and resemblance theories (theories arguing that music is expressive/meaningful in virtue of resemblances it shares with human expressiveness) tell a plausible story about why music is meaningful. Though they say little of what the systematicity of musical meaning inheres in. Worse, these theories often rely liberally on the role of the imagination in listening such that all of the questions we had about musical meaning could equally apply to imagination itself. If one argues that music is meaningful in virtue of our imagining a persona to whom heard emotional expressions are attributed, we still need an account of why we imagine what we ultimately imagine. These imaginative constraints are often largely omitted from such theories.

Separately, we also find what could be called a Hard Problem of Musical Meaning, to borrow from David Chalmers' terminology (Chalmers, 1995). We sometimes see explanations of musical meaning rooted in non-philosophical disciplines (e.g. neuroscience and cognitive science) as well as in philosophy (e.g. Formalist school of thought) that purport to explain musical meaning while appealing exclusively to essentially non-meaningful phenomena/elements. To put it crassly, many of these theories are some variation of: we notice that *x* pattern of neural activity/cognitive processes/formal description are co-occurrent with experiences of musical meaning, therefore, *x* must be responsible for musical meaning. Though patterns of neural activity, cognitive processes, or formal descriptions, in principle, could not explain musical meaning on their own for the same reason Chalmers contends consciousness could not be explained by appeal to patterns of neural activity and cognitive processes. Namely, such cognitive/neuronal patterns say nothing about why hearing musical stimuli involves *experience*. We can know everything there is to know about the neuronal/cognitive processes or formal descriptions associated with music without ever explaining why there is a *feeling of what it's like* to hear a piece of music.

GMG is one such theory in the cognitive domain. Though to the credit of Lerdahl and Jackendoff, they do not make claims about the implications the cognitive processes they posit have on affect. This makes their theory that much more amenable to unification with another that does make claims about musical affect, namely EMC. My hope is that in unifying EMC and GMG, we can arrive at an understanding of how the two components of musical experience interact, giving this theory of musical meaning a wider explanatory scope than those that have preceded it. Though it should be emphasized that the Analogical Argument is a continuation of previous investigations of musical meaning. While the past theories of musical meaning I've mentioned here fall short of adequately explaining musical meaning, they provide valuable insights into what this explanation must be like, and I rely heavily on these insights in order to develop the Analogical Argument.

I now begin by explaining EMC.

Section 3: Elementary Musical Meaning

In *Music and Embodied Cognition: Listening, Moving, Feeling, and Thinking,* Arnie Cox presents the *mimetic hypothesis* or EMC, a theory of how musical meaning is, in part, a product of mimetic cognition. 'Mimetic' refers broadly to imitation; mimetic cognition refers to the cognitive processes that mediate imitation. EMC provides an answer to our first question of why we experience affect when hearing musical sounds on the basis of two theses: (1) *we mimetically engage with different kinds of musical sounds (e.g. vocals, instruments), and* (2) *mimetic engagement entails the experience of affect.* Together, these theses entail (3) *we experience affect when hearing musical sounds.* 'Mimetic engagement' simply entails the activation of mimetic cognition in response to some stimuli. In this way, EMC provides a theory of how music could invoke affective states at all (addressing Q1).

I proceed by explaining each of EMC's two core theses, beginning with (1). Explaining this claim requires an explanation of how we could mimetically engage with sounds at all (let alone as reliably as EMC hypothesizes we mimetically engage with music). Explaining how we mimetically engage with sounds then requires explaining some fundamental forms and sub-processes of mimetics in general, (outside of musical or otherwise audial contexts), and as they pertain to EMC. I begin by establishing this conceptual foundation.

There are a variety of ways one might imitate something else. In humans, mimesis is perhaps most apparent in the interactions of infants and caregivers. These interactions are often characterized by the caregiver's repetition of gestures or facial expressions in hopes that the infant will imitate them. According to EMC, the infant's eventual imitation of these motor actions grants them a better understanding of "*What's it like to do that?*" and "*What's it like to be that?*" (Cox, 12). Through imitating these actions, the infant associates the performance of the action with the affective states it invokes in them and uses this as a basis for recognizing the affective states of others (Cox, 16). If the infant comes to associate pleasure with the action of smiling based on the co-occurrence of the two, they can recognize that their caregiver is experiencing pleasure when they recognize their smile—this ability to recognize the affective states of others on the basis of our observing them perform motor actions for which we have developed affective associations, as well as the notion that we typically do this, are basic assumptions of EMC. As the infant comes to draw stronger associations between motor actions and affective states, they begin to share the meaning of these actions with their caregivers and can perform the actions themselves with communicative intention.

This kind of overt imitation does not explain how we could mimetically engage with sounds, and is demonstrably not present in all musical experiences. We need a broader definition of imitation that accounts for covert imitation of sounds. Thus, Cox goes on to make a distinction between two forms of mimesis: Mimetic Motor Action (MMA) and Mimetic Motor Imagery (MMI)[4]. MMA refers to imitation through overt bodily movement, as exhibited in infant-caregiver interactions. Conversely, MMI is a form of covert mimesis where imitation occurs mentally as opposed to occurring through overt bodily movement. The 'motor imagery' in Mimetic Motor Imagery refers to the imagination of one's own actions/muscle movements. For example, one might imagine themself performing a dance move they've only just seen for the first time as an act of preparation for possibly performing the move themself. Cox argues that as we get older, we increasingly rely on MMI over MMA (Cox, 16). This would make practical sense in light of the fact that MMI is highly convenient—rather than overtly imitating every

[4] The existence of MMI is empirically supported by the experiments of Calvo-Merino et al. (2005) and Iacaboni et al. (2005), among several others, which demonstrate activation of brain areas associated with motor actions (call these motor areas) in the absence of overt motor movement after visual perception of others dancing or grasping a mug, respectively.

action we observe, we can covertly imagine ourselves performing these actions and achieve similar benefits (imagining ourselves performing a dance move gives us some useful indication of whether we could pull it off). Our imaginative capabilities increase with practice, and we can eventually mentally model more exceptional actions with greater accuracy, which then eases our reliance on MMA.

This foundation allows us to begin drawing connections between mimetics and sound (which will eventually allow us to draw connections to music). The first point to acknowledge is that all sounds are caused by physical events and are themselves evidence of the physical events that caused them. That is, sounds are non-arbitrarily related to their physical causes such that they provide clues about certain aspects of their physical causes. For example, a loud, booming sound is more indicative of a large or heavy physical cause than a small or light one. Certain aspects of sounds (loud) are non-arbitrarily related to certain aspects of their physical causes (heavy). This fact is presumably uncontroversial, but it gives way to the idea that we make inferences about the physical sources of sounds based on sounds themselves (which is again a basic assumption of EMC).

In making such inferences about the physical causes of sounds, we imagine what they must be like, and are consequently engaged in MMI. To clarify, 'inference' as it is used here refers to the kinds of inferences that are automatic and often unconscious. That is, we do not perform these inferences deliberately, and the results of these inferences, not the processes of inferences themselves, are overlaid on our experience. For example, we may hear (but not see) someone performing a tap-dancing routine (assume we have a concept of tap-dancing and some relevant experience with the actions/sounds associated with this act). The frequency, intensity, etc. of the sounds of the dancer's foot-taps serve as indications of the motor actions (at least certain aspects of them) that caused them. We might infer from the sounds of two nearly simultaneous and heavy foot-taps that the dancer has jumped. In making such inferences, we imagine the movements of the dancer, and are consequently engaged in MMI. Again, this process of inference occurs unconsciously and automatically such that we find the MMI (the image of the dancer jumping), but not the inference that led to it, in experience. Thus, we can mimetically engage with sounds caused by motor actions/physical events[5].

Although we mimetically engage with sounds, it seems clear to me that we generally mimetically engage with music much more extensively than with other non-musical sounds on the basis that musical experiences are so much more affectively rich than experiences of non-musical sounds. Cox accounts for this differential mimetic engagement by arguing that the inference of goal-directed physical causes is more likely to lead to mimetic engagement, and that we detect goal-directedness in music.

We can first consider the relationship between goal-directedness and mimetic engagement. The inferences we make about the physical causes of sounds may vary in the extent to which they indicate a goal-directed source. For example, we may hear (and not see) thudding against the steps of a stairway, the turning of a doorknob, and the creaking of an opening door. In such situations, we would likely infer that there is a person who has walked up some stairs and opened a door—a series of goal-directed actions. The notion that goal-directedness plays an important role in what we mimetically engage with is empirically supported by behavioral research on imitation indicating that we tend to focus more on the goal of an action rather than the action itself (Wohlschläger, Gattis, and Bekkering, 2003); other research similarly indicates

[5] This assertion is empirically supported by fMRI experiments demonstrating activation of motor areas when performing hand actions and when only hearing (and not seeing) others perform hand actions (Gazzola, Aziz-Zadeh, and Keysers, 2006). The hand actions used in the experiment were simple tasks such as ripping a piece of paper.

that MMI is more likely to be activated in observation, or inference, of more goal-directed actions (Grèzes et al., 1998), as determined by how much neuronal activation is observed in areas of the brain taken to be responsible for motor action/imagery.

Conversely, when we hear the sound of thunder, we typically don't think some goal-directed agent has caused it[6], as the sound of thunder reliably has merely physical (that is, non-goal-directed) causes. On the other hand, the goal-directedness of the causes of some sounds may be ambiguous. If you live in an apartment and hear a thump overhead, there would not be enough evidence to confidently infer what caused the sound (or the extent of its goal-directedness); perhaps your neighbor stomped their foot in frustration or a book happened to fall from their bookshelf. Such sounds are not as obviously indicative of the goal-directedness of their physical causes because we associate them with physical causes that are both goal-directed and not.

There is thus a continuum of mimetic engagement/inferred goal-directedness: on one end of this continuum are sounds such as those of walking up a flight of stairs or opening a door, the physical sources of which, according to EMC, we reliably model through MMA or MMI (per the behavioral research above) because we identify them as goal-directed actions. At the other end of this spectrum are sounds for which we would likely not infer a goal-directed physical cause.

[6] Granted, we can imagine cases where, for example, someone literally believes in the Zeus of Greek Mythology and infers the sound of thunder to have a goal-directed cause on this basis. However, these are generally outlier cases where our inferences are influenced by high-level psychological states (in this case, belief). For present purposes, we need not concern ourselves with such cases in order to explain how EMC explains the invocation of musical affect, as only considering cases that don't involve high-level psychological states are sufficient for telling the story EMC seeks to tell. Further, while it is interesting to consider what the limits of this influence are, as well as how other high-level psychological states such as attitudes or desires might exert an influence on inferred goal-directedness, we lack a theory of the manner and extent to which different high-level psychological states exert influence on musical meaning. These high-level psychological states include belief, such as a literal belief in Zeus, which may be the primary cause for the kind of inference thus described. The influence of high-level psychological states on musical meaning is discussed as an area for future research in Section 8.

Given that goal-directedness, to some extent, determines the extent to which we mimetically engage, we would not reliably mimetically engage with these sounds.

The sounds of music largely belong to the former end of the continuum. Music is unique in that we generally mimetically engage with music considerably more reliably than with non-musical audial stimuli. According to EMC, this is because we detect goal-directedness in music. We understand musical sounds to be reliably caused by goal-directed actions because we have all observed acts, however subtle, of musical performance, such as humming a tune. Inferring a goal-directed cause leads more reliably to mimetic engagement. We infer goal-directed causes for music to a greater extent/frequency than for non-musical sounds, and consequently mimetically engage with them to a greater extent/frequency than for non-musical sounds.

Thus far, we have become familiar with mimesis, explained how we mimetically engage with sound, and explained why we mimetically engage with music to a unique extent. At this point, we have essentially explained EMC's assertion that 'we mimetically engage when we hear different kinds of musical sounds'. However, one potential objection to EMC as it has been presented thus far is that if our capacity for mimetic engagement through MMI is grounded in our past motor experiences, then we should have great difficulty mimetically engaging with sounds produced with instruments/techniques that we ourselves have never performed with. Further, how can we mimetically engage with digital musical sounds, which have no obvious physical cause? Given that affect is reliably invoked by such sounds while their mimetic basis is not obvious, this would seem to put pressure on EMC.

We can answer these questions by introducing a distinction between intramodal and intermodal MMA/MMI, the latter of which allows us to mimetically engage with a target without having extensive prior experience/knowledge about it.

Intramodal, or direct-matching MMA/MMI refers to imitation where one tries to match the motor movements of another person by moving the same muscles as them in the same way (in intramodal MMA) or imagining moving the same muscles in the same way (intramodal MMI). Intramodal MMA can be observed in contexts where one mimics the playing of a musician using the same instrument; intramodal MMI can be observed in contexts where one imagines themself moving the same muscles as a dancer. Crucially, in intramodal mimetic engagement, we seek to imitate all aspects of a target's movement.

In contrast, intermodal, or cross-modal MMA/MMI refers to mimetic activity that imitates some, but not all, aspect(s) of its target. For example, we may engage in *acoustic* cross-modal imitation by singing or whistling the melody produced by an instrument. Such an imitation fails to be a direct-matching, because the physical arrangement of our vocal chords (which determines the sound that is produced) and that of any other instrument is simply different. It is in this sense that we represent some aspect—pitch in this example—of what we hear across modes, or 'cross-modally'. We are using different muscles and actions than those used by the original performer to mimic the original performance with respect to a particular aspect of it. Similarly, we may engage in *nonacoustic* cross-modal imitation by bobbing our head along to a rhythm. In these cases, we are imitating rhythmic aspects of heard sound.

Thus, our having a capacity for intermodal MMA/MMI explains how we can both engage with sounds produced with instruments/techniques that we ourselves have never performed with and engage with digital sounds. Despite perhaps not having the knowledge/experience necessary for intramodal MMA/MMI of musical sounds (mimicking all aspects of the motor actions that caused them), one can still model the music on a more abstract level through intermodal MMA/MMI (which only requires mimicking some aspects of the causal motor actions). In fact, we can engage in Intermodal MMA/MMI when only mimetically engaging with singular aspects of musical sounds such as their rhythmic or tonal profile (per the examples above). For example, we can tap our foot along to the rhythms of a drummer or sing a melody played by a trumpeter. While we may not all have extensive technical or theoretical knowledge of these instruments, we have basic musical intuitions for a variety of components of the music they produce, for example the rhythm and tonality (more on this in Section 5). This gives us enough to conceivably cross-modally mimetically engage with all musical sounds (which characteristically have rhythmic and tonal/harmonic components).

If we have an ability to mimetically engage with those musical sounds which we have no direct knowledge/experience how to produce on the basis of abstract musical features, we can apply this same ability to digital sounds which may have no existing acoustic musical instrument that could be played to make the sounds digital instruments make.

Having explained how we mimetically engage with music, we can now move on to the second thesis of EMC: mimetic engagement entails affect.

Recall the example of mimetic behavior in infants at the beginning of this section. The infant observes a motor action, mimics it, and comes to associate the affect that is invoked in them by performing the motor action with the action itself. This basic example can be used as a model for understanding how affect entailed by all MMA: the affect we experience in MMA is invoked by our performing (mimicking) a motor action, and we form motor action/affective state associations on this basis.

The associations we form between performed motor actions and invoked affective states can then guide the invocation of affective states through MMI. We may come to draw a strong association between smiling and the feeling of pleasure, say, through MMA. If this association is strong enough, Cox conjectures that MMI of observed smiles could invoke the feeling of pleasure in one's self. Imagining ourselves performing the action involves imagining the affect we would experience in performing the action, which consequently actually invokes the affect in us to varying extents. Imagine, for a moment, accidently sliding your finger across the edge of an envelope, resulting in the slow slicing of flesh. Did this imaginative exercise cause you to wince? While you might not have felt pain with the same intensity of actually cutting your finger (thankfully), it should have at least evoked a weaker/less vivid sense of pain (affective state). This is enough to render Cox's claim that MMI evokes affective states plausible.

So if both MMA and MMI involve the invocation of affect, and mimetic engagement just means one is engaged in MMA or MMI, then mimetic engagement involves the invocation of affective states. If this is true, and we mimetically engage with musical sounds, then we experience affect when listening to music. We have now answered Q1.

Section 4: Critical Evaluation of EMC

With this theoretical exposition behind us, we can now evaluate EMC's propositions as well as its role in the current project. Addressing these concerns will help us get clear on how EMC could complement GMG, and what its role will be in developing the Analogical Argument.

One primary shortcoming of EMC is its limited scope. Although it provides a robust explanation of how affect is invoked when listening to music, it seems to falter in explaining the specific influence of tonal/harmonic patterns and regularities on musical meaning. These patterns and regularities can be rhythmic or harmonic and observed in many musical patterns and regularities such as metrical structures, keys/scales, melodic variation, etc. To take one specific example, it's not clear how EMC could explain the particular musical affect associated with hearing the repetition of a chorus section of a composition. Explaining the role of such musical patterns and regularities in musical affect is important because, perhaps the key distinguishing feature between groups of pitch-time events that would standardly be labeled 'music' and those that wouldn't is some degree of organization (which is to say patterns and regularities) among pitch-time events.[7] Patterns and regularities thus seem fairly essential to what we would standardly consider to be music.

To be sure, EMC does provide some explanation of rhythmic affect—affect invoked in virtue of rhythmic patterns and regularities. We know what it is like to perform common rhythmic motor actions—for example, walking—and we also know what sounds are associated with the performance of these actions. When we hear sets of pitch-time events that have a similar rhythmic profile to the sounds we associate with walking, we attribute the affect associated with walking to the pitch-time events as an act of MMI. For example, a musical piece evoking a

[7] To further reinforce this intuition, we can notice how distasteful a musical composition devoid of any rhythmic or harmonic regularity sounds.

feeling of excitement in us may be attributable, in part, to its rhythmic profile resembling the rhythmic profile we would have in performing motor actions while excited (e.g. fast movements).

However, EMC fails to account for the affective states that are invoked in virtue of certain tonal/harmonic patterns and regularities. For example, major and minor chords, generally invoke characteristically different affective states (often described as 'happy' or 'sad', respectively). Though the difference in motor action that it takes to produce each of these chords on almost any chordal instrument would be about as minute as the difference between reaching for two adjacent keys on a computer keyboard. The core tenets of EMC would suggest that differences in the affect invoked by certain sounds should be proportional[8] to differences in their physical causes. Given this, it is unclear how EMC could explain this stark difference in the affect evoked on the basis of an unanalogously small difference in resembled motor action from the perspective of EMC. Meanwhile, we can tell some story about how virtually any musical rhythm acquires its affective attribution via resemblance with some rhythmic motor action or another.

Like harmonic affect, EMC also struggles to account for tonal affect, or the affect associated with tonality. EMC can account for some abstract tonal *forms*. For example, we may hear a saxophone raise its pitch throughout a phrase and liken this to a person's asking of a question, where they may manipulate the pitch of their voice in the same way. There are myriad resemblances we can identify between the pitches of instrumental phrases and speech acts: we

---

[8] That is, differences in musically invoked affective states should be proportional, in *some* sense, to a difference in associated motor action/imagery. There are different ways in which this 'proportional correspondence' could be understood. For example, Cox appeals to what he calls *exertion schema* when explaining why certain sounds invoke certain affective states; we might hear certain pitch-time events as being more intense than others in proportion to the extent to which we associate these sounds with motor actions that involve great exertion, for example. Despite there being different ways of understanding this 'proportional correspondence', there is no clear 'proportional correspondence' between differences in harmonic affect and associated motor actions/imagery.

can imagine a saxophone mimicking a groaning or screaming voice for example. However, this does not explain how the feeling arises in virtue of a particular tonal organization of pitch-time events. There are certain melodies that to some extent or another reliably invoke certain kinds of affective states, and EMC does not explain why this is so—certainly not as comprehensively as EMC's explanation of rhythmic affect. With EMC's account of rhythmic affect, we have an explanation of how the components of rhythm, namely the individual beats, can be arranged so as to resemble rhythmic motor actions. To have an analogously comprehensive account of tonal/harmonic affect, we would need some explanation of how certain arrangements of individual pitches could resemble motor actions (or the sounds they cause).

EMC may offer explanations of other sources of musical affect at an abstract level, but so long as it doesn't contend with the constitutive units of tonality/harmony (pitches), it fails to comprehensively account for the myriad feelings that can be evoked in virtue of certain tonal/harmonic organizations.

It is also worth noting EMC's reliance on neuroscientific evidence when assessing its claims. As with all neuroscientific evidence, we can not posit causal relationships among certain patterns of neural activity (motor area activation) and certain mental states (mimetic engagement). This is because all we can observe through the observation of neural activity is the co-occurence of certain neural activations. However, the higher the reliability of this co-occurrence, the more confidence we can have in certain patterns of neural activity and certain mental states predicting each other. Given that the co-occurence of motor area activation and overt motor actions/MMI is in fact highly reliable, we can be confident in the claims EMC makes on the basis of this relationship.

While these concerns certainly present problems for EMC, we should not let them overshadow the immense value that the mimetic hypothesis offers. EMC presents a plausible, scientific explanation of how music can invoke affect at all, distinguishing itself from numerous past theories that fell short of this goal. This constitutes an answer to Q1. It also makes some progress in explaining the question of how it could be that the affect that is invoked in musical experience is non-arbitrarily related to heard musical sounds (Q2). Though to fully answer this second question we will have to become acquainted with GMG, which I turn to now.

Section 5: Theoretical Exposition of GMG

EMC has provided an answer to how pitch-time events can have affective meaning at all, but there is more work to do to answer the question of why pitch-time events have the meaning that they do (Q2). What we want is an explanation of the systematicity of musical affect. I will argue that GMG's interaction with the musical meanings provided by EMC provide this explanation. This is the Analogical Argument alluded to in Section 1. The Analogical Argument, which will be defended in Section 6, will argue that GMG mediates our relation of pitch-time events so as to give rise to emergent musical meaning in essentially the same way (analogously) that the cognitive processes posited by Noam Chomsky in his Theory of Generative Linguistic Grammar (GLG) mediate our relation of terms so as to give rise to emergent linguistic meaning. More specifically, the Analogical Argument can be summarized as follows: GMG represents the musical analog for GLG's linguistic grammar. EMC represents the musical analog for semantics. Therefore, EMC and GMG facilitate the construal of musical meaning in a manner analogous to how GLG facilitates the construal of linguistic meaning.

Given that GLG has had immense success in explaining emergent linguistic meaning (and the systematicity of it), and that GMG employs the same Cognitivist approach as GLG, we ought to expect that GMG can help explain emergent musical meaning as well as the systematicity of it. Put simply, we should take GLG's success as a theory of emergent linguistic meaning to be indicative of GMG's success as a theory of emergent musical meaning on the basis that they both posit a grammar—a set of constraints on how we relate constitutive components of complex linguistic expressions (GLG) or complex combinations of pitch-time events (GMG).

With the goal of the Analogical Argument in mind, I now give a brief overview of GLG, so that we may be mindful of what it is we will be analogizing GMG to. I then proceed by explaining the core tenets of GMG and elaborate on reasons we should take GMG to be an equivalently useful theory for explaining the systematicity of musical meaning.

Briefly, GLG explains linguistic meaning by positing innate grammatical structures that explain key features of linguistic competency. One such grammatical structure is syntax, the cognitive process that mediates our understanding of language on the basis of the arrangement of terms. According to GLG, syntactical rules facilitate the systematicity and compositionality of language. 'Systematicity' in this context basically means that the meaning of linguistic expressions are determined by the meanings of their constituent terms, and the syntactical relations among them. For example, linguistic expressions such as 'the tiger runs' can not change in meaning unless their constituent terms or the syntactical relations among them change, as these determine the meaning of expression. 'Compositionality' in this context basically means that understanding a complex linguistic expression ensures our understanding of other complex linguistic expressions that have the same syntactic relations among terms, even when the meanings of the terms are different. For example, one might know the meaning of the expression 'the tiger runs fast' as well as the meaning of the term 'dog'. Even if we have somehow never encountered the expression 'the dog runs fast', we can know what this phrase means by virtue of compositionality. Additionally, semantics is the study of how linguistic terms have meanings—how we associate certain concepts with certain terms.

We're now ready to turn to GMG. Though before exploring the core tenets of GMG, I'd like to make some preliminary remarks about the theory that will be relevant for present purposes.

Firstly, the primary goal of GMG is an account of how we generate structural descriptions of music. The structural description is a mentally constructed formalization of musical stimuli and relations among them (not unlike deep structures in GLG, which are formalized descriptions of terms and relations among them). Lerdahl and Jackendoff take the structural description to represent our musical understanding, and their theory outlines the mechanics of the processes which lead to it.

Additionally, Lerdahl and Jackendoff are clear about only aspiring to give an account of those aspects of musical cognition that are hierarchical. 'Hierarchical' in this context roughly refers to our ability to make distinctions between local and global events/relative importance within certain aspects of music (and 'importance' in this context roughly means having an influence on how we hear other pitch-events). For example, we can create a hierarchy of rhythm in a piece by grouping certain sections on the basis of rhythm, and making some evaluation of which of these groupings are more influential to listening.

Non-hierarchical aspects of music include timbre, dynamics, and motivic-thematic processes (Lerdahl and Jackendoff, 9). These aspects cannot be easily organized hierarchically because, put simply, hierarchical organizations require that one instance of the aspect being organized could be considered subordinate to another instance. It is not clear what this would mean for timbre, dynamics, and motivic-thematic processes, because they resist division into levels or groupings of subordinate relations in a way rhythm, for example, does not.

While these non-hierarchical aspects of music are undoubtedly influential to musical meaning, we need not worry too much about their omission in GMG. This is because, the influence these aspects have on musical meaning could already be explained by EMC (certainly with regards to timbre and dynamics). For example, the meaning associated with changes in dynamics (which is roughly synonymous to volume) could be accounted for by our associating different affective states with louder or quieter sounds. Further, as will be outlined in Section 6, what we want from GMG is a story of how musical meaning is influenced by *relations* among pitch-time so as to give rise to emergent meaning. That these aspects can not be related (because they are not hierarchical) disqualifies them from contributing to this goal.

Next, Lerdahl and Jackendoff do not make claims about the implications the structural description or the processes that lead to it have for musical affect. This is because, in their words, "...it is hard to say anything systematic [about affect] beyond crude statements such as observing that loud and fast music tends to be exciting. To approach any of the subtleties of musical affect, we assume, requires a better understanding of musical structure" (Lerdahl and Jackendoff, 8). Having been familiarized with EMC, we may disagree with these assertions while at the same time recognizing how they already suggest a complementary relationship between EMC and GMG.

Finally, it is worth acknowledging how GMG's being a theory of musical intuitions renders its claims subject to tests of introspection. If there is some musical intuition concerning how we group pitch-time events cross time, then we (or well-trained musicians, at the very least) should be able to notice the role of this process in musical experience. This is not to say that we should expect all listeners to be able to verbally report the hierarchical grouping structures that they mentally generate while listening to music, but it is to say that the process of grouping should be distinguishable enough from separate processes for us to be able to confidently say that *some process along those lines* must exist.

To illustrate this notion, we can consider a linguistic example. When listening to spoken language we can introspectively observe different cognitive processes at work, and with effort, recognize the ways in which they are distinct. We might hear a child say, “ice cream, me want” and notice that, while we may be able to understand the meaning the child wishes to convey, their statement is somehow ill-formed (not well-formed). The error we detect is not necessarily about the meaning of the words used, but rather their arrangement in the sentence. This ought to indicate that there is some linguistic intuition having to do with sentence structure at work (namely syntax) that is distinguishable from another linguistic intuition that deals with the meaning of words (namely semantics). We need only consult these intuitions through introspection to be able to tell them apart. It is this kind of introspection that ought to be analogously applicable to musical intuitions.

We can see this ability exhibited in considering what might standardly be considered music. For those sequences of pitch-time events that would standardly be considered music (or a musical piece), surely none would have ten-minute gaps of silence in the middle or have melodies where every note is five octaves away from the previous note.[9] Such collections of pitch-time events would seem to defy intuitive constraints for music, and do so on the basis of organizational features. The “song” with a ten-minute period of silence may be universally labeled as music if we simply eliminate this period of silence—notably, all that would change is the temporal distribution of pitch-time events (rhythm). Similarly, the “song” with melodies where every note is four octaves away from the previous note could be universally labeled as music if we simply put all adjacent pitch-time in the same (or adjacent) octaves—notably, all that would change is the pitch distribution of the pitch-time events (tonality).

[9] The distance between two notes with respect to pitch is an interval. Put simply, an ‘octave’ is a relatively large interval. The intervallic relationships among adjacent pitch-time events in the vast majority of popular melodies (and melodies we would standardly label as music) are generally much smaller than an octave. It is on this basis that an interval of five octaves among adjacent pitch-time events is intuitively unmusical.

These cases highlight rhythmic and tonal intuitions, respectively. They demonstrate how our intuitions can guide our theorization of musical cognition; the rhythmic intuition just described suggests the necessity of something like the grouping structure Lerdahl and Jackendoff posit, and the harmonic intuition suggests the necessity of something like the prolongational reduction Lerdahl and Jackendoff posit, for example. While we might not definitively know what the processes are, we *can* know certain things about what they must be like. Thus, tests of introspection can provide support for GMG insofar as they align with the processes posited by Lerdahl and Jackendoff.[10] I turn now to the main tenets of GMG.

There are four main cognitive processes that Lerdahl and Jackendoff posit as generating our structural descriptions (which I will simply refer to as 'processes' moving forward): 1. grouping structure, 2. metrical structure, 3. time-span reduction, and 4. prolongational reduction. Additionally, GMG posits a distinction between *well-formedness rules* and *preference rules.* In the context of GLG, well-formedness refers to whether or not a linguistic expression adheres to grammatical rules. In the context of GLG, well-formedness rules specify the possible structural descriptions that a given piece can have, while preference rules determine which structural descriptions are ultimately adopted.

Such a distinction is needed because in music it is possible to generate different, but equally valid structural descriptions for any given piece, where 'valid' just means that the description abides by well-formedness rules. This is in contrast to language, where there is only one correct structural description, or deep-structure, for any given well-formed statement—this is

[10] There exists some empirical support for GMG, and experiments could in principle be designed to test any of the listener's intuitions about the structural components of music. This is because all of GMG's claims are about cognitive processes, and cognitive science has already developed robust scientific methods for the experimental testing of hypotheses about cognitive processes. Lerdahl and Jackendoff hypothesize that these structural descriptions can then be compared to the cognitive processes they lay out. The research that has been conducted in this regard generally supports the existence of the kinds of processes that Lerdahl and Jackendoff posit (Hansen, 2010).

because it is based on strict term-definition relationships[11]. Music is not strict in this way, but rather relationally defined (an argument that will be defended in Section 7). This allows for some variation in the kinds of structural descriptions that are valid, though the variation among these descriptions is first constrained by well-formedness rules, which specify the function of each process and the domains in which they're to be applied. In other words, while well-formedness rules determine the *kinds* of inferences to be performed on audial stimuli, preference rules determine the manner and extent to which they are applied. As an example, we can consider grouping structure's third well-formedness rule which says that 'A group may contain smaller groups' (see Figure 4 on the next page for illustration). This rule allows for smaller groupings of pitch-time events, but it is constrained by grouping structure's first preference-rule, which says to 'avoid analyses with very small groups—the smaller, the less preferable'. The well-formedness rule defines the *kind* of inference to be performed, while the preference-rule determines the extent and manner to which the well-formedness rule applies.

The preference rules in fact do most of the work in this theory, though they are omitted here. Instead, an understanding of the functions of the cognitive processes (well-formedness rules) themselves is most important for present purposes. This is because the Analogical Argument will rely on a musical analogy to GLG, and well-formedness rules are sufficient for supporting this analogy.

Briefly, I will now define each of the four processes GMG posits:

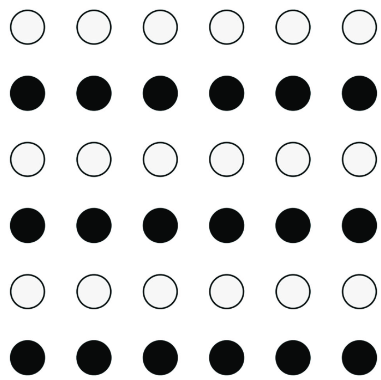
1. Grouping structure refers to, "...hierarchical segmentation of the piece into motives, phrases and sections" (Lerdahl and Jackendoff, 8). At a broad level, this component of musical

---

[11] Granted, in some cases it might be ambiguous what the correct structural description of a well-formed statement is when considered in isolation (e.g. "The boy saw the man with the telescope."). However, determining the correct structural description of such statements would be a matter of knowing the meaning that it was meant to convey. This is still in contrast to music where different structural descriptions could be considered valid even when there is agreement about what musical meaning was sought to be conveyed by a particular musical passage.

grammar is analogous to visual Gestalt psychology which posits that we group objects in our visual field on the basis of their similarity. Consider the following figure:

Figure 4



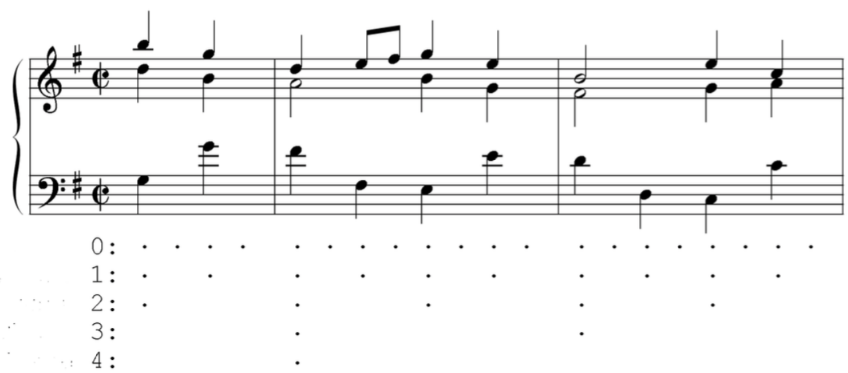
We tend to group the black dots with the other black dots along the same row, and white dots with other white dots along the same row, on the basis of similarity. In the same way we have this innate capacity to group similar objects in our visual field, we group pitch-time events into motives, phrases, and sections:

Figure 5

These groupings differ in terms of how local/global they are, though they all indicate some thematic cohesion (basically, a kind of similarity) among pitch-time events.

2. Metrical structure refers to, "...the intuition that the events of the piece are related to a regular alternation of strong and weak beats at a number of hierarchical levels" (Lerdahl and Jackendoff, 8). 'Strong' and 'weak' in this context refer to the extent to which a beat interacts, or occurs simultaneously, with other beats at different levels of hierarchical organization. Metrical structure is formalized in the following:

Figure 6

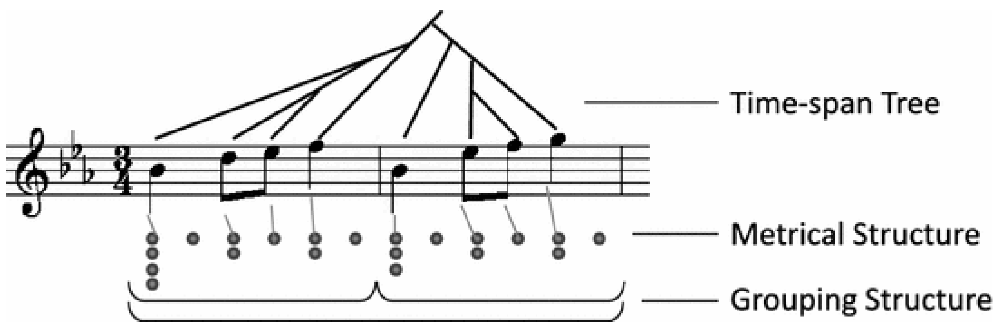


Put simply, the columns with more dots (beats) are 'stronger'. Metrical structure is related to grouping structure in that it indicates the temporal 'backdrop' on which musical groupings are situated. Importantly, this metrical structure is regular in that the lowest level of its hierarchy is composed of equally-spaced time segments, and all events at higher levels of the hierarchy align with some segment at the lowest level.

3. Time-span reduction, "...assigns to the pitches of the piece a hierarchy of 'structural importance' with respect to their position in grouping and metrical structure" (Lerdahl and Jackendoff, 8). To unpack this, metrical structure deals strictly with the temporal organization of pitch-time events whereas grouping deals more directly with their melodic/thematic cohesion. Both of these more basic processes assign a certain relative importance (again, 'importance'

indicates certain pitch-time events having implications for how we hear other pitch-time events) to pitch-time events, and their interaction gives way to a more abstract kind of relative importance—namely time-span reduction. This is illustrated in the following:

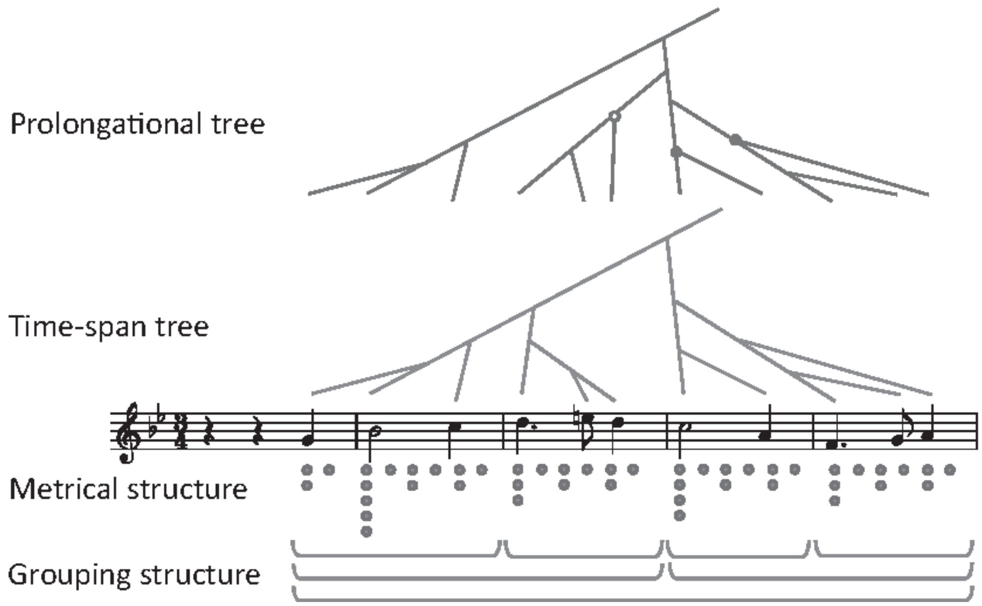
Figure 7



The role of grouping and metrical structure in determining time-spans can vary by how global the level of analysis is: at local levels, time-spans can be totally determined by metrical structure. On more global levels, time-spans can be totally determined by grouping structure. This variance in role occurs on a gradual basis.

4. Prolongational reduction, "...assigns to the pitches a hierarchy that expresses harmonic and melodic tension and relaxation" (Lerdahl and Jackendoff, 9). Prolongational reduction, as well as the rest of the processes thus described, are formalized in the following:

Figure 8



Prolongational reduction is difficult if not impossible to define on a purely formal basis because it is defined in terms of tension and relaxation, which become apparent in musical experience—they are things we *feel* in response to music (as opposed to merely representing, though certain representations in structural descriptions are strongly correlated with tension/relaxation). This difficulty is exacerbated by the fact that the tension and relaxation in a piece can be modified by tonality and/or rhythm. With respect to pitch, tension can be manipulated by the amount of dissonance among the pitches of different pitch-time events. Though there is not a direct correlation between pitch dissonance and tension; the kind of dissonance, as well as time the dissonance is introduced, are both influential in how overall tension is modified. Rhythmic tension can be increased with greater speed or rhythmic irregularity, for example, where irregularity roughly refers to a lack of overlap between grouping and metrical descriptions.

There is also not a direct correlation between speed/irregularity and tension; the tension is emergent from particular styles and placements of speed/irregularity in relation to the rest of the piece. This complexity is then further exacerbated by the interaction rhythmic tension has with tonal/harmonic tension. Again, GMG fails to account for why such relationships obtain because it is not concerned with affect. Though we take prolongational reduction to be highly indicative of the patterns of tension/relaxation that we would hear in musical experience because certain formal patterns are highly associated with certain affective states. Presumably, this is because these formal patterns, when realized in musical experience, reliably lead to the invocation of certain affective states via EMC. As previously mentioned, GMG's inability to account for musical affect here further demonstrates the potential contributions of the Analogical Argument—unifying GMG with EMC would provide a theoretical basis for explaining the correlational relationship certain formal descriptions of music have with the affect they evoke.

We are now familiar with both GMG and GLG. Again, the core proposal of the Analogical Argument is the idea that GMG could be analogized to GLG, such that GMG could explain how musical meaning emerges in the same way that GLG explains how linguistic meaning emerges. We are ready to begin substantiating this analogical relationship. I turn to this substantiation in Section 6.

Section 6: The Analogical Argument

Now that we are armed with knowledge of both EMC and GMG, we can begin exploring the ways they interact. Our main goal remains the explication of musical meaning, and the task of providing this explanation can be guided by the following variation on Q2: *if EMC basically says that the affect invoked by music is the affect we experience in motor action, then in virtue of what is musical experience not affectively cacophonous? That is, if all the different sounds that constitute a piece of music each invoke some distinct affective state, why do we seemingly experience aggregates of these affective states systematically and reliably in musical experience, and not a disordered conglomeration of them?*[12] In the remainder of this thesis, I demonstrate that this question can be answered by the Analogical Argument. This argument can be summarized as follows: GMG represents the musical analog for linguistic grammar. EMC represents the musical analog for semantics. Therefore, EMC and GMG facilitate the construal of musical meaning in a manner analogous to how GLG facilitates the construal of linguistic meaning.

The Analogical Argument will, of course, be centered around analogies between language and music. Though, these analogies have their limits, and it will eventually become easier to characterize musical meaning through *dis*analogies with language. These disanalogies will prominently include the relational nature of musical meaning, and the resistance music has to the kind of total incomprehensibility that can be found in ill-formed linguistic utterances. It is also worth mentioning that the Analogical Argument, while primarily emphasizing those aspects of language and music that are analogous, will be interspersed with caveats defining the exact extent and manner in which certain aspects of music and language can be analogized. One should

[12] For example, when we hear melodies, we don't experience the musical meaning (affective states) of the different pitch-time events in the melody separately, but rather hear the melody itself as having its own meaning. It is in this sense that we experience 'aggregates' of elemental musically invoked affective states.

not expect any of the analogies presented here to have perfect one-to-one correspondence across domains.

In presenting these ideas, I begin by clarifying the problem posed by the motivating questions presented at the beginning of this section. I then explain the Analogical Argument. Finally, I further characterize our understanding of musical meaning via disanalogies with language.

To fully grasp the weight of the problem posed by this section's motivating question of why we experience the aggregate affective states that we do in musical experience (and not a disordered conglomeration of distinct affective states), we can (unsurprisingly) rely on analogies to language: When crafting sentences, we arrange words in certain orders so as to convey a particular meaning. The meanings of sentences are, in part, determined by the meanings of the terms that constitute them. Though crucially, the meanings of sentences go beyond the meaning of terms themselves—there is an emergent meaning that is understood via a linguistic *grammar* (GLG).

We can consider the terms 'tiger', 'run', and 'fast'. Each of these terms have different meanings, and when they are arranged in a particular order so as to form a sentence (ex. 'tiger(s) run fast'), there is an emergent meaning that is not contained within any of the meanings of the individual terms. We can know this because while the sentence contains the meaning of each term, it also contains some other meaning. Namely, the meaning that *tigers run fast*. The fact that this meaning cannot be reduced to more basic components of the sentence without sacrificing something further reinforces the idea that the sentence carries a meaning of its own, distinct from the meaning of its constituent terms/concepts. It has an emergent meaning, determined by its constituent terms.

Notice that the emergent meaning of the sentence is also not contained in a mere agglomeration of terms. To confirm this, we can consider other such word orders:

Utterance	**Well-Formedness**
'Tiger fast run'	Ill-Formed
'Fast tiger(s) run'	Well-Formed
'Fast run tiger'	Ill-Formed
'Run tiger fast'	Ill-Formed
'Run fast tiger'	Could be well-formed if we charitably interpret this utterance as a command: 'Run fast, tiger'

Certain permutations of these terms give rise to an emergent meaning at the level of the sentence, and others don't. The question then is, *in virtue of what are some of these utterances not mere agglomerations of terms?* This is the same situation we currently find ourselves in with respect to music.

In language, we now know that the sentence's emergent meaning is attributable to a linguistic grammar, which GLG explains. The 'tigers run fast' exercise above serves to highlight the role of syntactic rules, posited by GLG, in linguistic meaning. Those sentences that are well-formed, and therefore have emergent sentential meaning, are those that abide by grammatical rules. We can now more pointedly rephrase our motivating question as *what musical analog could there be for linguistic grammar?*

With this, we can begin developing the Analogical Argument. Once more, this argument can be summarized as follows: GMG represents the musical analog for linguistic grammar (GLG). EMC represents the musical analog for semantics. Therefore, EMC and GMG facilitate the construal of musical meaning in a manner analogous to how GLG facilitates the construal of linguistic meaning. I now explain these assertions in turn, beginning with 'GMG represents the musical analog for linguistic grammar'.

Given that GLG deals in relations among terms, we ought to consider what the musical analog of these might be in order to properly analogize GMG and GLG. Put simply, terms are basic units of linguistic meaning. They are the smallest linguistic units which can be said to have distinct meanings. Moving forward, I'll refer to the smallest unit of meaning (in music or language) as a 'morpheme': the most basic element that can be said to carry a unique meaning. As mentioned, linguistic morphemes are terms. With regards to music, defining the morpheme is a much more difficult task. It could be argued that in a hierarchy of musical classification (ex. note, melody, section, composition), the note would have to constitute a morpheme on the basis that those with absolute pitch experience distinct feelings for certain notes: there's a feeling of *what it's like to hear a G-sharp* for those with absolute pitch that is distinct from the feelings invoked by hearing other notes. If the invocation of a distinct affective state is the criteria we are using to determine whether a musical sound has meaning (as we have done in discussions of EMC), then it seems as if individual notes would have to carry meaning.

However, absolute pitch is rare. The vast majority of listeners of music have relative pitch, which basically means that they don't recognize individual notes as invoking particularly distinct feelings. Rather, meaning is evoked by musical sounds in virtue of the relationships between them: while an isolated note might not evoke a distinct feeling, two notes heard simultaneously or in succession could evoke a feeling that is distinct to their intervallic relationship (how far the notes are from each other with respect to pitch). One may still contend that, even for those with relative pitch, isolated pitches can have affective meaning in virtue of subtle forms of MMI—in this case, distinct feelings would be evoked by our knowing something of what it would be like to produce the sound ourselves, not necessarily by the sound's pitch. This is a valid concern that is worth considering and serves to emphasize how difficult it can be to define a threshold for musical meaning. Though to illustrate why isolated pitches would not constitute musical morphemes, we can consider those we encounter in everyday life; we do not standardly consider the sound of a train horn or the ringing of a digital alarm clock to be 'music', nor would we say these sounds invoke affective states to a significant extent. On the other hand, we could imagine how the same sounds *would* be considered 'music' if they were varied in pitch and/or had some rhythmic organization.

With these complexities in mind, our best route forward is to abandon the goal of finding a precise definition for musical morphemes. Rather, we can rely on our shared intuitions about what juncture in a hierarchy of musical classification evokes musical meaning that meets some (undefined) threshold level of salience/vividness, and this will be sufficient for our purposes. This threshold could be satisfied by a short melody or a single chord for example. While it may very well be the case that we construe meaning from individual notes on the basis of MMI, it's not clear that this meaning would be a particularly salient or vivid one; on the other hand, most would agree that short melodies or individual chords would have a salient/vivid meaning such that these sounds would be widely regarded as musical.

We can now return to the present claim that GMG represents the musical analog for GLG. I will unpack this claim by considering how all of the cognitive processes posited by GMG (grouping structure, metrical structure, time-span reduction, and prolongational reduction) might map on to this analogy. Thus, in order to substantiate the claim that GMG represents the musical analog for GLG, I will argue that: Sub-sentential grammatical structures can be roughly analogized to GMG's grouping structure and metrical structure, sentential grammatical structures (as well as more complex grammatical structures) can be roughly analogized to GMG's time-span and prolongational reduction, and culturally- or context-dependent applications of syntax can be roughly analogized to GMG's preference rules. I now explain these assertions in turn.

The claim that sub-sentential grammatical structures can be roughly analogized to GMG's grouping structure and metrical structure is made on the basis that sub-sentential grammatical structures, and grouping and metrical structure, fulfill analogous roles in the construal of meaning. Firstly, when speaking of 'sub-sentential grammatical structures', I refer to grammatical structures such as verb phrases, noun phrases, prepositional phrases, etc.; these are 'sub-sentential' in the sense that they could not independently constitute a well-formed sentence.

Sub-sentential grammatical structures fulfill analogous roles to grouping and metrical structure in that, while they can contribute to emergent meaning, they are not sufficient for constituting what we can call a 'complete thought' or 'complete expression'. For example, we can consider the noun phrase, 'the majestic tiger', and notice that it intuitively seems to be missing something ("What *about* the majestic tiger?"). Similarly, we can consider the verb phrase, 'is running quickly'. Considered in isolation, this verb phrase, a sub-sentential grammatical structure, also intuitively seems to be missing something ("*What, or who,* is running quickly?").

This sense that something is missing from these statements is the intuition that a 'complete thought' has not been expressed (moving forward, I'll refer to these as 'complete expressions', as this term is more clearly applicable to music). This is because complete expressions generally require both a subject, and some idea about, or involving, the subject. For example, ‘the majestic tiger is running quickly’, has both of these attributes: the majestic tiger is the subject and ‘running quickly’ is some idea involving the subject. This is therefore a complete expression.

That grouping and metrical structures are both insufficient for constituting complete expressions is suggested by the fact that they could just as easily accurately represent myriad different musical passages, all of which would have distinct meanings. This is not unlike how we could imagine the noun/verb phrases presented above as belonging to myriad different well-formed sentences (e.g. ‘The majestic tiger gazed menacingly’, ‘The majestic tiger passionately plays the harp’). To highlight this point, we can simply observe how two pieces evoke different feelings when they differ only in tonality. For example, we can consider ‘Happy Birthday’ in a major[13] and minor[14] key. Because grouping and metrical structure deal only in rhythmic relationships in a piece, and these different versions of ‘Happy Birthday’ differ only with respect to the pitches of their constitutive pitch-time events[15], we can be confident that either version of this piece shares exactly the same grouping and metrical structure. Nonetheless, they invoke different feelings. This is analogous to how sub-sentential grammatical structures could subsist myriad different emergent sentential meanings without constituting a complete expression on their own. It is in this sense that grouping and metrical structure can be analogized to sub-sentential grammatical structures.

Sentential (and more complex) grammatical structures can be roughly analogized to GMG’s prolongational reduction and time-span reduction. This claim is made on the basis that both sentential grammatical structures and prolongational/time-span reduction fulfill analogous

[13] Happy Birthday - Major Key
[14] Happy Birthday - Minor Key
[15] These pieces are in major and minor keys of the same root, which basically means that the difference in the pitches of the pitch-times constituting the piece is minimal, and despite this we can observe a marked difference in the affect they invoke.

roles in facilitating the emergent meaning of complete expressions. With regards to language, we can support this claim by recognizing that the reason ‘the majestic tiger gazed menacingly’ is a complete expression is because sentential grammatical structures relate the meanings of these sub-sentential grammatical structures to each other.

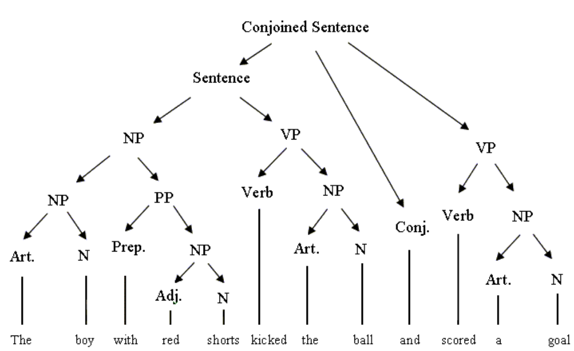
To outline this particular analogy, we can recall the following: time-span reduction, “...assigns to the pitches of the piece a hierarchy of ‘structural importance’ with respect to their position in grouping and metrical structure” (Lerdahl and Jackendoff, 8). Again, importance in this context roughly refers to the extent to which a group of pitch-time events has influence over how we hear other pitch-time events. Also, prolongational reduction, “...assigns to the pitches a hierarchy that expresses harmonic and melodic tension and relaxation” (Lerdahl and Jackendoff, 9). While time-span reduction and prolongational reduction are more closely associated with rhythm and tonality respectively, they each assign hierarchies of structural importance to music.

These processes are analogous to sentential grammatical structures in the sense that they determine how we relate those aspects of musical structural organization that subsist complete expressions; more specifically, they determine how we relate grouping and metrical structure, which don’t give way to emergent meaning on their own, so as to give way to emergent meaning. We can think of sentential grammatical structures as organizing and relating sub-sentential grammatical structures on the basis of their function within a sentence. By assigning terms complementary functional roles within a sentence (ex. subject-verb-object), we can know how the meanings of the individual terms ought to be related so as to give rise to an emergent sentential meaning**.** As a basic example, we can consider ‘Tigers run’. ‘Tigers’ and ‘run’ have particular meanings when considered in isolation, though in identifying each term as the subject and verb within a sentence respectively, we are led to interpret the terms such that we can construe emergent sentential meaning.

Similarly in music, we can think of Lerdahl and Jackendoff's concept of structural 'importance' as a kind of syntactical function that certain pitch-time events can instantiate. Indeed, our working definition of importance just *is* the property of certain pitch-time events influencing how we understand other pitch-time events—not unlike the way identifying 'tigers' as a subject influences our understanding of 'run' as something the subject (tigers) does. Contrast this with the meaning of 'run' considered in isolation: something that things (not necessarily the relevant subject, tigers) do. The specification of the function of certain pitch-time events would then determine how we relate the meanings of different pitch-time events with each other in complementary ways, and consequently give rise to an emergent meaning that is not inherent in any of the individual musical morphemes, or grouping and metrical structures.

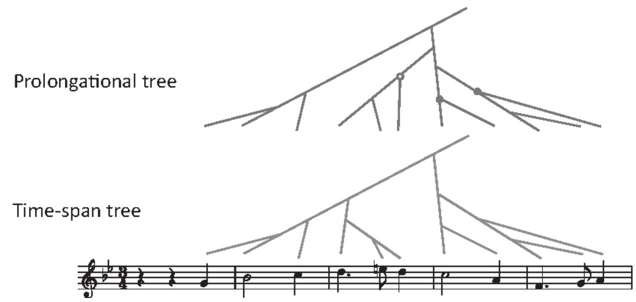
Indeed, this analogous relationship between syntax and time-span/prolongational reduction can be superficially appreciated in noticing how similarly they are formalized:

Figure 9



Syntactical Structure

Figure 10



Prolongational/Time-Span Reduction

GMG's preference rules can be roughly and broadly analogized to more context- and culturally-dependent applications of syntax. I begin with the musical component of this claim. Recall that preference rules are not cognitive processes in and of themselves. Rather, they determine the manner and extent to which well-formedness rules (which essentially just *are* the four cognitive processes constituting GMG) are applied to heard musical stimuli. Importantly, Lerdahl and Jackendoff claim that well-formedness rules are universal among listeners, while preference rules can vary on the basis of an individual's previous exposure to music—a unique history of musical experiences can lead to unique proclivities with regards to the structural description. For example, listeners from different cultures have been observed to prefer metrical structures with different metrical levels when presented with the same musical stimuli (Cameron, et. al, 2015).

Similarly, context- and culturally-dependent applications of syntax are not universal. It is in virtue of our having had musical experiences particular to a culture or context that we can understand the meaning of such idiosyncratic applications of syntactic rules.

However, we ought to keep in mind that GMG cannot be responsible for all idiosyncratic interpretations of musical meaning—the affect that is invoked by mimetic engagement is also subject to variation through varied experience. In some cases, this variation might be such that two different listeners characterize a passage using descriptors with categorically different meanings (e.g. 'gleeful' and 'triumphant'). This is in contrast to language where highly disparate experiences may lead to slightly varied understanding of terms, but often still lead to the same essential meanings. With this, we can now consider the role of EMC in the Analogical Argument.

EMC can be roughly analogized to GLG's semantics. Again, semantics is the study of linguistic meaning, as opposed to the other cognitive processes which all either mediate the construal of meaning or take meaning as their input. Semantics is how meaning is introduced to our language processing at all. Similarly, EMC deals in affective meaning directly through previously described processes of mimetic engagement. This is in contrast to GMG, which deals in organizational representations of musical stimuli. Again, Lerdahl and Jackendoff make clear that GMG is not meant to be an explanation of musical affect. However, once affect (meaning) has been introduced by mimetic engagement with musical sounds, the structural description generated by GMG relates these sounds and their associated affective meaning so as to give rise to an emergent meaning not inherent in any of the individual morphemes.

To review, in language we have terms/concepts and relations among them; linguistic grammar is the mechanism by which these relational structures are competently ascertained as well as generated so as to give rise to emergent sentential meanings. In music, we have sounds/affective meanings and relations among them. EMC explains the process by which we arrive at these associations between sounds and affective meanings. GMG, broadly, is the mechanism by which these relational structures are competently ascertained as well as generated so as to give rise to an emergent musical meaning.

This is generally as far as analogies to language will get us in understanding musical meaning. Though disanalogies to language prove similarly useful.

The first disanalogy to recognize is that musical sounds don't have veridicality conditions in the way linguistic utterances do—this is to say that they can't be 'wrong' in the common sense of the term. For example, the expression 'Tammy the tiger runs fast' could be true or false on the basis that its meaning is in accordance with certain facts about the world. If Tammy the tiger in fact runs fast, then this statement is true. If Tammy the tiger in fact runs slow, this statement is false. Linguistic utterances have veridicality conditions because terms represent things in the world. Terms are *about* things, and by extension, so are sentences. If what a statement is saying about some thing, state of affairs, etc. is not in accordance with the facts of the matter, then the statement is false.

If EMC is correct, then music has no such veridicality conditions because musical sounds do not refer to or represent things in the external world—they aren't *about* things. Rather, musical sounds simply invoke affective states *we* have or have had in performing motor actions. Where terms refer to external things, musical sounds invoke internal things. This invocation of affective states resists veridicality because we can't be wrong or right when we feel. Feelings themselves aren't about the world in the same way terms are. If we stub our toe against a hard surface, in what sense can we be wrong about when we feel pain?[16] We know when linguistic statements are wrong by investigating facts about the world. Though with regard to affective

[16] To be sure, there are cases where we may be 'wrong' in our belief that 'I am in pain'. For example, we could consider psychosomatic conditions such as phantom limb syndrome. However, even in these cases, it is true that people experience pain; pain is a part of their phenomenology as if it in fact did have physiological causes. This is the sense of 'wrong' we are interested in.

states, we can only 'investigate' through introspection/awareness of the contents of our consciousness.

However, our interpretation of musical meaning can be 'wrong' in the sense that we could generate a structural description that is in fact an incorrect application of well-formedness rules to a particular musical stimulus. GMG's well-formedness rules have sufficient conditions that can either be satisfied or not satisfied by musical stimuli. For example, we can consider metrical structure's first well-formedness rule: *Every attack point must be associated with a beat at the smallest metrical level present at that point in the piece*. Therefore, if our structural description is such that there is an attack point that lies between two segments of our metrical structure, then we have generated an incorrect structural description. Further, there can be a fact of the matter as to whether distal musical stimuli satisfy such conditions because they are based on objective features of pitch-time events. With regards to metrical structure's first well-formedness rule, an attack point is the 'front end' of a sound—it refers to the moment at which a sound begins. We can objectively determine the occurrence of an attack point by identifying when a sound begins following a period of silence (through using a microphone to measure volume levels, for example). All of the well-formedness rules postulated by GMG are subject to such tests of satisfaction. Therefore, we can objectively determine when an incorrect structural description has been generated.

This brings us to our next disanalogy: while an incorrect application of linguistic grammar rules (e.g. 'fast run tiger') fails to result in an emergent meaning, incorrect structural descriptions can still give rise to emergent meaning. To understand how this is so, we must first recognize the relational nature of musical meaning. Musical morphemes may have a particular meaning when heard in isolation, but this meaning can be significantly augmented by placing this morpheme in different musical contexts. For example, we can consider 'the lick'—a melody constituted by seven notes that is often used in jazz music, among other genres.[17] Depending on the context, the musical meaning of 'the lick', and of musical morphemes more generally, can be of a categorical difference: the same melody can sound happy or sad depending on the context.

Contrast this with linguistic terms whose meaning is stable regardless of their context. 'Tigers' refers to tigers whether we say 'tigers run fast', 'tigers have found a cure for cancer', 'this song is an ode to tigers', etc. Granted, there may be instances where a term's meaning is varied. For example, 'she's a tiger on the trade floor' employs a metaphorical use of the term 'tiger'. Nonetheless, such uses of terms fundamentally rely on the stable, *essential* meaning that the term has in isolation; it is only in virtue of the word's stable, essential meaning that it has the metaphorical meaning that it does in certain contexts.

This serves to demonstrate how musical morphemes can have implications on each other's meaning. To further illustrate this possibly counterintuitive notion, we can consider analogies to vision. Gestalt Psychology has identified principles of visual perception such as good continuation and completion. When these principles are applied to visual perceptions, such as patterns of dots, we see the placement of some dots as having implications for the placement of others (Koopman and Davies, 2001).

We can consider the following figure as a visual situation in which we find such Gestalt principles in action:

[17] 'The Lick' in different musical contexts.

Figure 11

While there is nothing in the arrangement of these dots explicitly indicating that they fall along the edge of a circle, we perceive them as having a circular orientation nonetheless. One of the groups of three dots in this figure may seem to display a pattern of being oriented around an arc. We have an expectation that other dots will continue this pattern, and therefore see them as such, as opposed to simply perceiving the dots as dots and not inferring a pattern to their positioning at all. It is in this sense that the placement of the dots have implications for each other.

So too in music, past pitch-time events have implications for how we hear future pitch-time events: possible patterns and regularities may predispose us to want to hear that pattern or regularity be continued. This relational nature of musical meaning then explains how we can have emergent meaning from a set of pitch-time events even when our structural description is wrong. Even when our structural description is wrong, it still specifies *some* relation among heard pitch-time events. All that is needed to construe emergent musical meaning is morphemes (EMC) and relations among them (GMG). Linguistic meaning may be characterized similarly, but crucially, we must recall: syntax requires that words be able to fulfill certain complementary functions so as to contribute to a complete expression, among other things. There is no emergent meaning that comes from a sentence exclusively composed of nouns, for example: ‘tiger tiger tiger’. This is because, at least very generally, sentences require a subject and some idea about or involving them. We can have complete expressions when the terms that constitute a sentence fulfill complementary functional roles. For example, in ‘tigers run fast’ these words fulfill the syntactic functions of subject-verb-adverb (in the order that they appear).

This is in contrast to music, where there is no such division of functional roles. The musical analog of syntax—time-span and prolongational reduction—deals only in one function. Namely, importance. Because there is only one function in ‘musical syntax’, it can be fulfilled by any musical morpheme. Therefore, so long as a structural description specifies *some* hierarchical relationship among pitch-time events on the basis of importance, then emergent musical meaning can be construed, regardless of whether or not the structural description is in fact in accordance with heard musical stimuli.

Section 7: Towards a Theory of Musical Understanding

Thus far, what has been presented is a theory of musical meaning; a theory of why music invokes affective states, and why it invokes the affective states that it does. We can now explore some notable implications the theory would have if it is true.

I'd like to begin this elaboration by considering the strength of a possible analogical relationship between musical meaning and linguistic meaning that has hitherto been left unexplored—namely, the extent to which each is universally accessible. Now that we have a story about what musical meaning is, we can inquire into the extent to which it is shared.

In language, determining whether some linguistic meaning has been understood by others[18] can be fairly straightforward. Often you can tell whether your expression has been understood based on whether a fellow interlocutor's expression or behaviors are consistent with the meaning of what you've said. If they respond with something completely unrelated to or inconsistent with what you've said, they've likely misunderstood you. There are instances in which it's unclear whether we've been understood: a conversation might get cut off before we get a response, for example. Though even in such cases, we can expect that knowing whether we've been understood would just be a matter of asking someone to respond to or rephrase what we've said, and seeing whether this corresponds with the meaning of what we've said.

Determining whether a musical meaning has been understood is a more complex and difficult task. Firstly, we should recognize a disanalogy between what it means to 'understand' a linguistic expression and what it would mean to 'understand' a musical expression, if 'understanding' can be appropriately applied to musical meaning at all. In language,

---

[18] By 'others' here, I refer to other humans. There exist AI chatbots that can generate appropriate linguistic responses to statements, but it's not clear they understand these responses in the same way humans do.

understanding should result in the same conceptual understanding in the minds of both the expresser and interpreter.

For music, insofar as it is a communicative apparatus, successful communication can not be a matter of achieving the same conceptual understanding because musical meaning is not essentially conceptual, but rather essentially affective. What we seem to mean when we say that a musical meaning has been understood is that the same, or sufficiently similar, affective states have been invoked in the minds of both the expresser and interpreter, or at least among interpreters.[19] Admittedly, the sense of 'similar' I use here is a broad one. We want to allow for varying affective responses to music to constitute a musical understanding because requiring identical affective states would be too chauvinistic; any difference in invoked affective states among listeners would disqualify the presence of an understanding. Worse, even if there somehow was a musical understanding in this chauvinistic sense, we would never be able to confirm that this was the case because it is impossible to definitively know the affective character of another's mental state. That is, we can't definitively know the affective nature of their subjective experience in the same way we know the nature of our own subjective experience (Nagel, 1974).

In addition to not requiring that affective states be identical among listeners in order for musical understanding to be achieved, we also don't want to allow for individually relativistic affective responses to constitute an understanding. If the affective states of listeners shared no meaningful resemblance at all, this would go against the very definition of 'understanding'. Though, we can have confidence that the musical affective responses of different listeners are not individually relativistic for the same reason we don't worry about the absence of similarities

[19] There exist artificial intelligences that can generate musical compositions which are basically indistinguishable compositions authored by humans (Cope, 2013), but we wouldn't say that these AIs understand the music they generate in the relevant (affective) sense.

among the mental states of others at large. Being unable to definitively know the character of the mental states of others, we can still have confidence that our approximations of others' mental states are accurate with some significant degree of fidelity; while we might not know whether someone is experiencing the exact same sadness (for example) we've experienced, we can know that they feel some kind of sadness along these lines such that we are justified in forming beliefs, claims, decisions, etc. on the basis of their verbal reports, bodily gestures, etc. Call this level of confidence, which grounds safe inferences of another's mental state, approximate knowledge of another's mental states. Contrast this with what we will call definitive knowledge of another's mental states; a precise knowledge of what it would be like to experience another's mental state.

Definitive knowledge of another's mental states can not be attained, though approximate knowledge can be. Thus, we can say that a musical understanding has been achieved when one could justifiably have approximate knowledge that the musically invoked affective states among listeners resemble each other. In this way, we can avoid both chauvinistic and overly relativistic requirements for musical understanding.

Again, as has hopefully been demonstrated, defining musical understanding in rigorous terms is a complex and difficult task. There is much more that could be said about what a musical understanding must consist in. Though as we have done for musical morphemes, we can avoid becoming overly concerned with particularities by simply relying on our shared intuitions about what a musical understanding is. This shared intuition can be observed in how musical experience is often characterized as fostering a strong sense of common meaning. We might feel a strong sense of connection with others while listening to music in communal settings, or feel that a piece of music has somehow expressed one's own feelings. This intuition can be quite salient, and widely shared[20], despite our not having an explanation of what our sense of musical understanding consists in. This said, our goal in this section will be to establish approximate knowledge of the musically invoked affective states of others. More specifically, our goal will be to identify a theoretical justification for approximate knowledge that musical experiences among individuals resemble each other, consequently justifying the intuition we often have that a musical understanding has been achieved. I will argue that the Analogical Argument as it has been presented thus far can indeed provide a basis for claims about the universal accessibility of musical meaning, though demonstrating this will require further elaboration of the Analogical Argument's claims.

This investigation is imperative because music is a cultural universal and a common part of everyday life for countless people. Considering this, having an ambiguous understanding of the universal accessibility of musical meaning constitutes a significant gap in our understanding of human experience at large. We should want to be able to confirm whether the apparent correspondence in our musical experiences is justified or not. Thus, accounting for the universal accessibility of musical meaning is a matter of philosophical importance.

In what is to follow, I argue that we can find theoretical justification for the intuitive universal accessibility of musical meaning by relying primarily on the core principles of EMC. In particular, the central role the mimetic hypothesis places on motor actions. To crudely restate the core claim of EMC, the affective states we experience while listening to music are a kind of 'fusion', as it were, of the basic affective states we come to experience and know through motor action. Thus, I will argue that we can have confidence that musical meaning is universally accessible as follows: if musical affect is based in part on motor affect (via EMC), and we have

[20] This intuition may not necessarily be salient and widely shared for the same musical stimuli, but is salient and widely shared with respect to different musical stimuli/experiences for different people.

good reason to believe that there is a significant degree of universality for motor affect, then we have good reason to believe that there is a significant degree of universality for musical affect insofar as it is based on motor affect. Note that this argument contends only with the semantic component of the Analogical Argument (EMC), though we should keep in mind the syntactic component (GMG) as well.

The first premise in this argument (that musical affect is based on motor affect) has already been defended on the basis of the mimetic hypothesis, and the conclusion follows logically from the premises. Demonstrating the truth of this argument will then be a matter of substantiating the second premise: the claim that there is a significant degree of universality for motor affect.

We should now consider reasons for believing there is a significant degree of universality for motor affect. Two reasons will be emphasized: the significant biological similarities among humans, and the similarity in motor action conceptualization that the mimetic hypothesis describes.

That there are significant biological similarities among humans is presumably uncontroversial. We now know that 99.9% of genetics are shared among humans (Lander, E. S. et al., 2001), for example, and many physiological similarities can be appreciated with the naked eye. Though if these biological similarities are to support claims of similarities in motor affect, the connection between biology and affective states must be drawn. I now defend the connection between biological states and affective states.

The relationship between minds and bodies has long been a topic of philosophical contention. While I acknowledge the importance of questions about whether and to what extent the mind is essentially distinct from the body, I want to avoid getting involved in such debates here. Rather, I'd like to highlight how, independent of what the fact of the matter is with regards to mind/body distinctions, we can draw safe conclusions about the mind and/or body on the basis of reliable relationships between nominally mental and physical phenomena. For example, physical behaviors such as verbal or gestural expressions can all serve to indicate a mental state. We find these behaviors to be reliably associated with certain affective states, and this allows these behaviors to provide a basis for approximate knowledge of another's mental state. Presumably even Dualists who believe the mind and body are distinct rely on such indicators to navigate the world; and justifiably so, as these relationships are reliable.

Further, if we *are* inclined towards the belief that the mind is emergent from physical phenomena, then we might be satisfied with a recognition of how, if we humans share 99.9% of our DNA, it seems like we can generally expect that the processes that give rise to experiences (neurology, physiology, etc.) are similar enough among people to realize similar mental states.

Establishing a connection between biological and mental states only helps us in establishing a significant degree of universality of motor affect among humans if we think that people in fact perform many of the same motor actions. I defend this assertion now.

Not unlike paradigm cases of mimesis, commonalities in performed motor actions are especially apparent in infants. As infants, our earliest motor experiences are reflexive: we consistently display the same set of basic and involuntary motor actions in response to certain stimuli. These include reflexes such as sucking, which is observed when an object is inserted into an infant's mouth, or the Moro Reflex in which infants spread out their arms upon detecting a loss in balance, among others (Modrell, A.K. & Tadi, Prasanna, 2021). Also, infants have been observed to instinctively step as if they were walking when held upright with their feet touching a surface (Yang, 1998), a motor action that might not seem so basic! The fact that such actions are observed in infants prior to their learning basic skills essential for navigating the world (e.g. talking) lends credence to the idea that these actions are biologically innate.

If there is such a thing as innately prescribed motor actions, then we have principled grounds on which to speculate that at least some basic motor actions, and their associated motor affect, are universally known and experienced by humans. Though we may worry that, while such empirically supported similarities in the motor action of infants are significant, they are not sufficient for establishing similarities in performed motor actions at large; that is, motor actions performed by non-infant humans.

As we get older and rely less on involuntary reflexes, our motor actions become more diverse. To be sure, many motor reflexes such as recoiling your hand after touching a dangerously hot object endure throughout our lives, and these secure at least some degree of similarity in performed motor action. Though their prevalence is limited to relatively exceptional circumstances, and our reliance on them is de-emphasized as we develop the capacities to navigate the world and make decisions about which motor actions we want to perform. This is in contrast to infants, who still lack such capacities, and therefore rely heavily on instinctive actions for their survival. There is comparatively little diversity in the motor actions of infants, and much of it can be summarized in terms of basic motor instincts.

In noticing this diversity of motor action among adults, we may even speculate that there are some motor actions that are so rare as to only be experienced by very few individuals and/or in very rare situations. Consequently, this would result in the motor affect associated with such actions being scarcely experienced. For example, we might think that the affect associated with walking on the moon simply bears no meaningful resemblance to the affect associated with walking on Earth, because walking on the moon involves negotiating gravity in a totally different way. All motor experiences are distinct in at least some sense, but perhaps we think that the affect associated with walking on the moon is distinct from the affect associated with more standard motor experiences, such that this 'distinctness' is of a kind difference from the 'distinctness' that the affective states of standard motor experiences have with respect to each other. If this is true, then it would seem to imply that the select few who have walked on the moon, among other rare motor actions, would experience different musical affect from others by way of EMC.

We can address this worry by appealing to core tenets of the mimetic hypothesis. Let us recall Cox's distinction between MMA (Mimetic Motor Action), which is concerned with overt action, and MMI (Mimetic Motor Imagery), which is concerned with imagined motor action. In Section 3, this distinction served to address a question similar to the one we face now: if musical affect is based on the affect associated with producing a certain sound/playing a particular instrument, why should we think someone who lacks knowledge of how to produce a certain sound/play a particular instrument is able to associate the right affect to heard musical stimuli? According to the mimetic hypothesis, we can bridge the gap between differential motor experiences by appealing to our capacity to simulate the experience of performing novel motor actions on the basis of our past motor experiences.

So while I might not know how to play the cello, I can get a pretty good sense of what it must be like through MMI based on my knowledge of what it's like to perform motor actions that resemble the motor actions involved in playing the cello (e.g. swaying your arm about the shoulder, pressing down on things with your fingers, etc.) or knowledge of what it's like to perform motor actions that cause sounds that resemble the sounds caused by a cello (e.g. subvocalization).

In principle, we should be able to imagine any motor action, musical and otherwise, we have never performed, and imagine the affect that would be associated with it in exactly the same way we imagine experiences of playing instruments we've never played. Indeed, playing an instrument is just a particular kind of motor action, and there doesn't seem to be a theoretical basis for concluding that the principles explaining our conceptualization of novel musical motor actions don't apply to motor actions more generally.

We can now relate this back to the problem at hand. The capacity for MMI as just described provides support for the notion that there is a significant degree of universality for motor affect because our past motor experiences, which constrain MMI, are similar enough to give way to similar MMI.

It is worth re-emphasizing here how remarkably similar our earliest motor experiences are. If we all share these motor experiences, as well as experiences of instinctive action in adulthood, then it seems as if when we utilize MMI, we are drawing on the same 'affective palette', as it were. While we may ultimately perform different motor actions, the fact that we share experiences of basic motor actions allows for our MMI to overlap significantly because they are formed with the same general constraints.

Relatedly, we can question whether motor actions beyond infancy are really as diverse as they might initially seem. While the motor actions of non-infant humans do not share the same level of resemblance as the motor actions of infants, they will usually bear some significant level of resemblance to basic motor actions such as jumping, outstretching your arms, stepping, clenching your fist, etc.[21] We can call such basic motor actions 'motor archetypes'.

---

[21] We might think this proposal can be alternatively defended on an Evolutionary basis: in the infinite range of possible human movement, having heuristics for determining which of these infinite possible movements is ultimately performed would be adaptive, and therefore selected for.

To illustrate this point, we can again consider what might be the most rare motor actions we know: Surely there is some especially unique affect associated with walking on the moon, and I would have to walk on the moon myself to really know what this experience is like. Though it seems as if I should be able to approximate the nature of this experience by relying on past experiences of, say, trying to 'walk' underwater; through imagination, we can sometimes use what we have previously experienced to achieve an understanding of what a novel experience would be like with some significant degree of fidelity (Kind, 2020). I might not know exactly what it's like to negotiate gravity on the moon, but I know what it's like to be in situations resembling walking on the moon, like being underwater: while underwater, a great deal of exertion is required to make small, slow movements and there is at least a feeling of negotiating a weaker gravitational pull than is normal. Being able to rely on this experience, as well as experiences of performing motor archetypes and possibly those of weightlessness while in a descending elevator, it seems as if I should be able to approximate what it's like to walk on the moon with some degree of fidelity over and above a guess. Granted, there would undoubtedly be many psychological factors influencing what the experience of walking on the moon is like, but we only need to account for the motor affect as it is this that is most relevant to EMC.

Even if it *were* the case that there is something about the experience of walking on the moon that is truly unattainable through MMI, we can still think this would not have great implications for musical affect because this act of walking on the moon would be one rare motor experience among many. Even for the astronaut who has been to the moon several times, there is good reason to believe that there would still be extensive similarity between their motor experiences and those of others (who have not been to the moon) such that there is still a significant degree of correspondence between their musical affect and that of others.

If what has been said is true, and we have good reason to believe that there is a significant degree of universality for motor affect, then we have good reason to believe there is a significant degree of universality for musical affect insofar as it depends on motor affect, via EMC. This argument goes a long way towards defending the universal accessibility of musical meaning. Though let us remember that musical meaning, according to the Analogical Argument, is determined not only by our basic affective responses to music (EMC), but also our structural descriptions of it (GMG). Having explained why we should think our basic affective responses to music are similar enough to subsist some level of universal musical meaning, we can now consider how GMG might figure in this discussion. I will argue that the implications GMG has for the universal accessibility of musical meaning are twofold: Firstly, the preference rules in GMG that allow for different, but equally valid, musical structural descriptions constitute one avenue by which musical meanings might vary among listeners. Secondly, GMG can help explain why affective invocations diverge amone listeners with varying levels of exposure to particular musical idioms, where 'idiom' refers to particular styles or genres of music (e.g. blues, heavy metal, classical etc.); the answer involves how the process of generating structural descriptions of music unfolds in real time, and explaining it will require further theoretical resources. I will turn to this problem of divergent affective responses to particular musical idioms after explaining the role preference rules play in the universal accessibility of musical meaning. I turn to this explanation now.

Let us recall that, as the Analogical Argument states, the emergent musical meaning we ultimately experience is determined by our basic affective responses to musical stimuli, but also how we relate these basic affective responses to each other—these relations are formalized in musical structural descriptions. Let us also recall that, in contrast to linguistic structural descriptions, Lerdahl and Jackendoff claim that musical structural descriptions involve preference rules. This basically means that there can be different, but equally valid, musical structural descriptions on the basis of preference/past musical experience.

Different relations among basic affective responses to music would necessarily entail different emergent musical meanings via the Analogical Argument. However, we ought to keep in mind that preference rules only allow for so much variance in the structural descriptions that could be considered valid. Adherence to well-formedness rules is the primary constraint on what could be considered a valid structural description, and these do not vary among individuals as preference rules do. Insofar as varying applications of preference rules lead to varying structural descriptions, agreement in the application of well-formedness rules ensures fundamentally similar structural descriptions. This said, preference rules do not pose a significant threat to the universal accessibility of musical meaning.

Thus, if we have principled reason to believe in the universal accessibility of musical meaning on the basis of EMC as well as GMG, and these theories constitute the core components of Analogical Argument (a theory of musical meaning), then we have principled reason to believe in a universally accessible musical meaning.

With this, let us evaluate what has been established. What has been presented constitutes an explanation of how we could justify the universal accessibility of musical meaning, *in theory*. Though we often find this theoretical correspondence among musical experiences to not be realized in practice; we can observe how verbal reports of the musical experiences of different listeners sometimes do not correspond with each other, for example. In particular, this occurs when one listens to music from an idiom with which they are not familiar. Avid listeners of jazz music may find great pleasure in a particular jazz performance, while the same music may sound unpleasant to someone who has had little to no exposure to jazz.

The Analogical Argument as it has been presented thus far does not provide an explanation for what goes wrong in such cases. It provides an explanation for how musical understanding could be achieved in theory; the resemblance among the affective states invoked in different listeners by the same piece of music should be a function of how much their motor experiences resemble each other (EMC), as well as how much their preference rules resemble each other (GMG). This remains true, though more theorizing will need to be done to explain why we find cases of individuals with fairly similar motor experiences and preference rules, but who have highly divergent affective responses to a given piece of music (as in the example of a jazz performance given above). Call such cases, where we find listeners who should have the same musical affective responses in theory, but don't with respect to particular musical idioms, cases of idiomatic divergence. I'd like to briefly propose a possible answer to this problem of idiomatic divergence, though doing so will require some exposition of Huron's ITPRA Theory (Huron, 2006), a theory about expectation as a source of musical affect, and explaining how this might interact with GMG and figure in the Analogical Argument's theoretical framework more broadly.

Before continuing, one might worry that if ITPRA Theory is needed to explain idiomatic divergence, then the Analogical Argument is not a complete theory of musical meaning. Though ITPRA Theory is merely an enhancement to the explanatory power of the Analogical Argument. ITPRA Theory does not contradict any of the claims of the Analogical Argument, and is compatible with the view that musical affect is basically the result of how structural descriptions organize elementary musical meaning so as to give way to emergent musical meaning. Though, as has been mentioned before, Lerdahl and Jackendoff theorize about the generation of structural descriptions of music from a time-independent perspective; they seek to explain only the 'finished product' of this process of generation. It goes without saying that the generation of structural descriptions unfolds across time. Thus, there is room for theorization about how affective responses to music might vary in real-time in relation to real-time generation of structural descriptions. ITPRA Theory provides such theoretical resources, and incorporating it increases the Analogical Argument's explanatory scope by allowing it to explain idiomatic divergence.

There are many philosophically interesting claims Huron makes in *Sweet Anticipation: Music and the Psychology of Expectation*, a book where he presents his ITPRA theory. Contending with all of them, as well as giving a comprehensive account of how the processes Huron describes interact with EMC and GMG at a cognitive level, is beyond the scope of the current project. I'd instead like to focus on the supporting role ITPRA Theory can play for the Analogical Argument; this can be achieved by maintaining focus on a few core concepts from Huron's ITPRA Theory.

Briefly, the core proposal of Huron's ITPRA[22] Theory is the idea that the affect that is invoked in musical experience is, in part, a result of certain predictive and evaluative proclivities we have as humans. We have certain expectations, informed by statistical exposure, about what impending events will be like (predictive), and make certain judgements or appraisals about the events we've expected once they occur (evaluative). These predictive and evaluative processes have affective components. To give some examples, we may feel tension or excitement prior to

[22] ITPRA stands for Imagination, Tension, Prediction, Reaction, Appraisal—these are the five response systems that Huron takes to be responsible for expectation-related musical affect

experiencing a particular event, or we may feel surprise or relief after having experienced an expected event.

I'd like to focus on a theme in Huron's account of the affective systematicity of ITPRA processes that will be crucial for our purposes: this is the idea that the character of the affective responses Huron describes are determined by the nature of our expectation. For example, in anticipation, whether our expectation is positively or negatively valenced can be informed by whether the thing we expect is a thing we desire to exist or occur. In evaluation, the character of our affective responses could depend on how well our expectation accorded with the relevant event. Huron theorizes that these expectational and evaluative proclivities are, in part, responsible for the affective states we experience in musical experience. Finally, not unlike Cox's claims about EMC, Huron takes ITPRA processes to be automatic and rooted in our physiology—roughly, Huron's claim is that expectations involve changes in our action readiness or behavioral proclivities. This has implications for our bodily states, which then has implications for affect (per the previous discussion on the connection between bodily states and affective states).

We can now begin integrating these concepts into the Analogical Argument. I argue that our musical expectations are primarily guided by our real-time structural description of music, a claim that will go against Huron's own view that our musical expectations are guided primarily by statistical exposure to certain sequences of sounds. To be clear, this is not to say that statistical exposure plays no role in our musical expectations. My claim is that the nature of our musical expectations can not *just* be a matter of the sounds and sound sequences we have had prior exposure to; our musical expectations are first constrained by certain rules of 'musical grammar', and the effects of statistical exposure on our expectations follow this. To briefly explain this disagreement, I take this position because proposing that statistical exposure exclusively guides our musical expectations leaves the question of why musical sounds/sound sequences have greater affective implications than other non-musical sounds/sound sequences for which we have had similar statistical exposure. Why don't we have vivid and salient affective states when hearing the sounds of clock tower bells, the sounds of household appliances like microwaves and washing machines, etc.? Further, if statistical exposure was all that was responsible for how we formed our musical expectations, why would it be the case that we often observe infants (who have presumably had much less exposure to music) reacting to music in ways that indicate similar affective responses as adults?

As has hopefully been demonstrated, the Analogical Argument can address such concerns. It does so, in part, by appealing to the process of structural description generation. Further, recall the discussion in Section 6 on the relational nature of musical meaning. Here, we had already identified an explanation of how structural description can guide our musical expectations: while generating structural descriptions, we notice certain patterns and regularities, and this predisposes us to expect these patterns and regularities to be completed or continued, or not. For these reasons, I propose that structural features of music, and not our statistical exposure to certain sounds, better serve as guides to our musical predictions.

Let us now return to the problem of idiomatic divergence. Musical idioms, again, are particular styles or genres of music. Though we can alternatively define musical idioms as sets of tendencies towards certain musical patterns and regularities. Jazz, for example, can be very loosely characterized as a style of music in which musical patterns and regularities such as ii-V-I chordal cadences and 'swung' rhythms, are prominent. This is not to say that if a particular piece of music does not include a ii-V-I cadence that it cannot belong to the Jazz idiom. Jazz merely has a tendency towards this particular structural pattern such that this tendency is characteristic of it. We say that a piece belongs to a musical idiom *x* when it conforms to the patterns and regularities that *x* has a tendency towards to some significant extent.

When we are familiar with a particular musical idiom, we are familiar with the structural patterns and regularities that are characteristic of it, and this knowledge can give way to accurate predictions about which patterns and regularities will be continued or not in a piece of music, and how this will happen. While listening to a jazz performance, if I hear a ii-V chordal cadence, and I know that the ii-V-I is a chordal cadence that is characteristic of Jazz on the basis of statistical exposure, then I can form the justified expectation that the ii-V cadence will continue to a ii-V-I cadence.

Contrast this with instances in which one is exposed to music from an idiom with which they are not familiar. In these cases, the listener does not have the knowledge that would be required to make justified predictions about how musical structural patterns and regularities will be continued or not. Consequently, a great deal of their predictions would just be incorrect.

To be sure, listeners often make incorrect predictions about music that belongs to an idiom with which they are familiar. To a certain extent, this should be desired, as Huron highlights the importance of surprise (caused by incorrect predictions) as a source of positively valenced musical affect. When listening to music from an idiom *x,* a listener who is familiar with *x* is more likely to have a pleasurable experience because, in addition to making some incorrect predictions about structural patterns and regularities, they will also make correct predictions. Again, this is in contrast to listeners who have not had exposure to *x* and are therefore more likely to make incorrect predictions than correct ones.

We can appreciate how this balance between correct and incorrect predictions is critical for pleasurable musical experiences when we consider Huron's concept of contrastive valence (Huron, 2006). Basically, this is the idea that the valence of a particular affective state is dependent on the character or valence of the affective states that precede it. For example, there is positively valenced affect associated with drinking a cool glass of water. Though this affect is especially positively valenced if we've spent a good deal of time outside on a scorching hot day. The contrast that the positively valenced affect of drinking a cool glass of water has with the negatively valenced affect of being outside too long on a scorching hot day serves to highlight and amplify the positive valence of drinking the cool glass of water. This is contrastive valence.

Huron takes contrastive valence to play a role in musical experience in essentially the same way it has just been explained. When listening to music, we have certain expectations about what will happen next. These expectations have physiological implications by way of priming action readiness or behavioral proclivities, and these physiological implications entail affective implications. When our musical predictions are inaccurate, there is a contrast between the affect that is associated with this, and the affect that is associated with accurate predictions we have made. Thus, having what would typically be considered a pleasurable musical experience involves a balance between accurate and inaccurate musical predictions. Further, this is why idiomatic divergence occurs. The listener that has had prior exposure to Jazz can have a pleasurable listening experience when listening to Jazz because they are able to make enough correct predictions about the music to achieve contrastive valence.

If this is true, we might be left wondering why, when we listen to highly predictable music (e.g. our favorite song), we don't always have unpleasurable musical experiences. In theory, it seems like such cases should be unpleasurable for the same reason listening to music that is highly unpredictable would be unpleasant—namely, there is no contrast between accurate and inaccurate predictions, and therefore no contrastive valence. As Nussbuam points out, the potential for music to invoke affective states even when we know what will happen in the music can be explained by appealing to the role of MMI (Nussbaum, 2007).[23] Even when we hear a piece of music for which we have already generated a structural description, we still imaginatively perform the actions that we associate with musical sounds; put another way, musical experience is something we *do*, as opposed to passively observe.[24] This is not unlike how, when recalling a memory involving the performance of motor actions (e.g. a memory of swimming underwater), we feel some of the affect we originally experienced in virtue of our mentally simulating past events.

[23] Nussbaum appeals to mental simulation, and not MMI explicitly, though the mental simulation Nussbau has in mind is essentially what Cox describes when he defines MMI.

[24] This doesn't exactly translate to musical experiences with too many incorrect predictions, because these incorrect predictions entail a disordered real-time generation of a musical structural description. Per the Analogical Argument, this would mean a disordered real-time relation among basic motor affective states, but also the motor imagery that coincides with them. Having disordered relations among our motor imagery is tantamount to not being able to have the continuous motor simulation (not being able to '*do'* music) that the other kinds of musical experiences in question (those with no contrastive valence as a result of correct prediction) have, and which secure ensure pleasurable musical experiences.

Section 8: Next Steps

In this thesis, I have proposed a theory of musical meaning, as well as a theory of how how and why musical meaning might vary among listeners. While I take to these theories to provide substantive contributions to philosophical and scientific investigations of music, it must be admitted that they are coarse theories, as it were. They explain musical phenomena at an abstract level, and a great deal more work would need to be done to achieve a comprehensive account of how we go from *particular* musical stimuli to *particular* affective states, if such a thing is even possible. We should be clear about what is still left to explain. These remaining questions can be roughly divided into those that involve clarifying the processes and capacities already considered in this thesis, and those that involve processes and capacities not considered in this thesis. I provide a brief overview of these questions here.

Firstly, as discussed in Section 4, both EMC and GMG fall short of explaining the role of harmony in musical meaning. By extension, the unified theory presented here also fails to explain the role of harmony in musical meaning, and this is problematic because harmony is involved in the vast majority of musical pieces.

Throughout this thesis, I have appealed to innate capacities playing a role in musical experience. Though there remains much room to clarify the precise extent to which such musical capacities are innate or not. Also, there is more left to learn about the neural and cognitive cognitive bases of these capacities; these investigations may lend support to the Analogical Argument or not. There has been some evidence indicating a neural correspondence between music and language processing; for example, Broca's area and the pars orbitalis region, both being brain areas associated with language processing, can be observed to be activated when listening to music (Tan et al., 2010). Additionally, suffering lesions to brain areas associated with music processing have been observed to have implications for language processing, such as not being able to discriminate between statements and questions (e.g. ‘That’s a tiger’ or ‘That’s a tiger?’) which differ as a result of pitch alteration (Tan et al., 2010). This said, there is also evidence indicating some neural independence of language- and music-processing.

We don’t yet know what role attention plays in this process. We know that music generally invokes affective states. Though this is apparently contradicted by our not having salient affective states when, for example, a piece of music is playing quietly and we are too engrossed in a conversation to engage with it. We may be able to register the musical stimuli we receive, yet not have strong reactions to them such that we have a musical experience, but don’t construe musical meaning (Section 2). This suggests our attending to the music plays a role in whether we perceive it as meaningful, but the mechanics of this process (e.g. How much attention is required? Does it matter which aspects of the music we pay particular attention to?) remain unknown.

Beyond such matters, we can inquire into the roles that high-level psychological states such as beliefs, attitudes, and desires play in musical experience[25]. What about musical meaning changes, if anything, when I believe the composition I’m listening to is about love, World War I, tigers, etc.? Also, intra-individual variations in musical experience provoke questions of the role that moods might play in our musical preferences. It is often unclear exactly how our moods affect our listening preferences (why do we choose the actual song/genre that we do?) though we can recognize a change in our preferences clearly enough for us to know that they are playing a role. Similarly, why do some people simply like or enjoy music more than others? We sometimes develop new tastes in music fairly spontaneously (that is, not as a result of differentiated

[25] This question is more germane to lyrical or program music, which does have an object, and which was not considered in this thesis.

statistical exposure or from a radical change in moods), and this is also mysterious. There also remain questions of how musical experience might vary based on the social context in which it is had. How does musical experience vary according to the gender, socioeconomic status, race, etc. of the listener?

Finally, we should look towards ways of bridging explanations of musical experience with aesthetics. It's clear that, aside from feelings, we sometimes find beauty in musical experience, and we can sometimes appreciate aesthetic value in music we don't necessarily enjoy. There does seem to be an intimate relationship between the musically invoked affective states that are explained by the Analogical Argument and perceptions of beauty, though the complex nature of this relationship has yet to be elucidated. We should strive towards moving beyond investigations of music as a mere conceptual and scientific artifact, and strive towards investigating music as an art form that can only be truly known through experience, and which has the power to add meaning to our lives.

Bibliography

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: an FMRI study with expert dancers. *Cerebral cortex (New York, N.Y. : 1991)*, *15*(8), 1243–1249. https://doi.org/10.1093/cercor/bhi007

Cameron DJ, Bentley J and Grahn JA (2015) Cross-cultural influences on rhythm processing: reproduction, discrimination, and beat tapping. Front. Psychol. 6:366. doi: 10.3389/fpsyg.2015.00366

Chalmers, David (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies* 2 (3):200-19.

Chomsky, N. (1966). Topics in the theory of generative grammar (Ser. Janua linguarum. series minor, nr. 56). Mouton.

Cope, D., & Silva, P. (2013). David Cope and Experiments in Musical Intelligence.

Cox, A. (2017). *Music and Embodied Cognition: Listening, Moving, Feeling, and Thinking*. Indiana University Press.

Dissanayake, Ellen (2000). *Art and Intimacy*. University of Washington Press.

Fard, A. (n.d.). *Gestalt Perception: Law of Similarity*. adamfard. Retrieved from https://adamfard.com/blog/law-of-similarity.

Gazzola, V., Aziz-Zadeh, L., & Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Current biology : CB*, *16*(18), 1824–1829. https://doi.org/10.1016/j.cub.2006.07.072

Grèzes, J., Costes, N., & Decety, J. (1998). Top-down effect of strategy on the perception of human biological motion: A PET investigation. *Cognitive Neuropsychology, 15*(6-8), 553–582. https://doi.org/10.1080/026432998381023

Harris, A. (n.d.). *Linguistic Deep Structure*. Pinterest. Retrieved from https://www.pinterest.co.uk/pin/562105597221632275/.

Iacoboni M, Molnar-Szakacs I, Gallese V, Buccino G, Mazziotta JC, Rizzolatti G (2005) Grasping the Intentions of Others with One's Own Mirror Neuron System. PLoS Biol 3(3): e79. https://doi.org/10.1371/journal.pbio.0030079

Johnston, A. (n.d.). *The Lick*. Quora. Retrieved from https://www.quora.com/What-is-the-lick-in-jazz.

Kind, Amy (2020). What Imagination Teaches. In John Schwenkler & Enoch Lambert (eds.), *Becoming Someone New: Essays on Transformative Experience, Choice, and Change.*

Koopman, C. and Davies, S. (2001), Musical Meaning in a Broader Perspective. The Journal of Aesthetics and Art Criticism, 59: 261-273. https://doi.org/10.1111/1540-6245.00024

Lander, E. S. et al., International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. Nature, 409(6822), 860–921. https://doi.org/10.1038/35057062

Lerdahl, F., & Jackendoff, R. (2017). *A Generative Theory of Tonal Music*. Cambridge: The MIT Press.

Metrical Structure. (n.d.). Music Practice and Theory. Retrieved from https://music.stackexchange.com/questions/44001/metrical-structure-levels-of-metrical-grouping.

Modrell AK, Tadi P. Primitive Reflexes. [Updated 2021 Mar 21]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK554606/

Nagel, T. (1974). What Is It Like to Be a Bat? *The Philosophical Review*, *83*(4), 435–450. https://doi.org/10.2307/2183914

Nietzsche, F. (2008). *The birth of tragedy*. Oxford University Press.

Nussbaum, C. O. (2007). *The Musical Representation: Meaning, Ontology, and Emotion*. The MIT Press. http://www.jstor.org/stable/j.ctt5hhjqv

Ordoñana, J. A., & Laucirica, A. (2010). Lerdahl and Jackendoff's grouping structure rules in the performance of a Hindemith sonata. *The Spanish journal of psychology*, *13*(1), 101–111. https://doi.org/10.1017/s113874160000370x

Quarter Notes. (n.d.). Kaitlin Bove. Retrieved from https://kaitlinbove.com/principles-of-notation.

Scott Rickard. (n.d.). *The world's ugliest music*. Retrieved from https://www.youtube.com/watch?v=RENk9PK06AQ&t=464s&ab_channel=TEDxTalks.

Tan, S.-L., Pfordresher, P., & Harré Rom. (2010). Psychology of music : from sound to significance. Psychology Press. Retrieved April 23, 2022, from INSERT-MISSING-URL.

Volume and Pitch. (n.d.). Flute & Guitar . Retrieved from https://vlinmp3.weebly.com/harmonics-and-timbre.html.

Wohlschläger, A., Gattis, M., & Bekkering, H. (2003). Action generation and action perception in imitation: an instance of the ideomotor principle. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *358*(1431), 501–515. https://doi.org/10.1098/rstb.2002.1257

Yaysuz. (n.d.). *The Lick Compilation Vii*. Retrieved from https://www.youtube.com/watch?v=hlY5fh3auR4&ab_channel=Yaysuz.