Perception, Expectation, Affect, Analysis

Elizabeth Hellmuth Margulis

ABSTRACT: Descriptive music analysis often aims to explain musical experience in terms of the characteristics of musical structure. But musical experience is largely subjective and varies from listener to listener. This paper uses a case study of expectation theory to demonstrate how sensitive analysis can account for commonalities in musical experience while reserving space for individual differences. The account relies on previous empirical work establishing that topical context can modulate the kind of affect generated by a syntactic surprise. According to this model, surprise generates perceived intensity broadly across experienced listeners, but the topical context helps select the dimension along which the intensity gets perceived. By teasing apart these mechanisms, a clearer delineation of which aspects of experience might be shared across listeners and which are more susceptible to individual differences can be established.

It is impossible for an analysis to reference expectation without invoking some listener. This is ironic given the grounds for Leonard B. Meyer’s initial enthusiasm for expectation: its supposed potential for eliminating precisely the complexity and subjectivity that listeners bring to the analytic enterprise. By Meyer’s reckoning, granted listeners who have developed reaction patterns appropriate to the work in question, the structure of the affective response to a piece of music can be studied by examining the music itself. […] the study and analysis of the affective content of a particular work […] can be made without continual and explicit reference to the responses of the listener or critic. That is, subjective content can be discussed objectively.1

1 Meyer 1956, 32.
Certain listeners, in other words, possess sufficient awareness of the typical patterns in the relevant style and share a core set of predictions so reliably that the analyst can simply assume them, and tie structural occurrence to affective consequence without worrying about that go-between, the listener. By directly linking structure to affect, this theoretical stance bypasses the problems of subjectivity and individual variation that typically arise when listeners are considered. Listeners who might be assumed to fulfill these criteria have come to be known as “experienced listeners,” an idealization most clearly summarized by Fred Lerdahl and Ray Jackendoff.\(^2\)

This assumption possesses all the alluring potential and subtle danger of any reductive take on music analysis, and the history of expectation’s use in music analysis has mostly been the history of more and less reflective usages of the construct of “the listener.” Just who the listener might be is a critical question on which fundamental disciplinary perspectives hinge. For a committed ethnomusicologist, for example, any particular listener might be so deeply shaped by the unique culture and set of experiences surrounding her that it is futile to think about “the listener” in any abstract sense. For the typical experimental psychologist, however, whose methodologies typically involve the identification of invariances among responses in a large group of listeners, the commonalities among listeners might be the central topic of interest.

Can these stances be integrated?\(^3\) A moderate stance might hold that despite individual variation, certain hardwired perceptual tendencies characterize responses almost universally. For example, as chronicled by David Huron, the genetically endowed startle response ensures that a sudden \textit{sforzando} will consistently elicit predictable effects even among listeners with very different backgrounds.\(^4\)

But restricting analysis to perceptual tendencies at the level of reflexes seems unsatisfactory; the percentage of musical experiences that we really care about that can be understood in terms of these very basic perceptual processes is probably small. If beyond reflexes lies intractable subjectivity, analysis is ill-suited to description and explanation, capable only of prescription (or what Temperley more gently calls “suggestion”\(^5\))—of outlining ways to hear a passage that this or that individual finds satisfactory. While the prescriptive enterprise certainly has its place in analysis, I am reluctant to give up on the enterprise of description: a type of analysis that can take a musical experience a listener might already be having and expose some of the forces and rationales that underlie it. Rather than propose new ways of listening, descriptive analysis illuminates modes of listening people already enjoy.

This paper claims that there is a space between universal, automatic responses and contingent, subjective experience within which analysis can do descriptive work. Expectation theory, while generally considered a branch of music psychology, positions itself within this space, seeking to explore the ways empirically verifiable psychological tendencies might result in rich, affect-full musical experiences that admit of variability.

---

\(^2\) Lerdahl/Jackendoff 1983, 3f.

\(^3\) See, for example, Korsyn 2003.

\(^4\) Huron 2006, 62.

\(^5\) Temperley 1999.
from person to person. In the end, it may still achieve Meyer’s goal of discussing subjective content objectively, but not by jettisoning the listener—rather by committing to sufficient specificity about the listener, and about the limits of what is shared among different listeners.

Mechanisms Of Musical Affect

In a target article in Behavioral and Brain Sciences, Patrik N. Juslin and Daniel Västfjäll provide what to my knowledge is the first comprehensive overview of the mechanisms by which music might evoke emotions.\(^6\) They propose six possible mechanisms: (1) brain stem reflexes such as the startle response; (2) evaluative conditioning, where music comes to be associated with some extramusical stimulus because they have coincided in multiple or significant ways; (3) emotional contagion, where internal mimicking of the music’s expressive properties comes to trigger feelings; (4) visual imagery that arises in connection with the music and itself possesses emotional valence; (5) episodic memory, where music triggers recollection of some life event that occurred while previously listening to it; and (6) expectancy, where music artfully violates and capitulates to the expectations listeners hold about what might come next. Of these six, musical structure in the sense that a theorist might think about structure only plays a significant role in a few. Brain stem reflexes relate to events (sudden \textit{sforzandi}, for example) that hardly require analysis to understand. Visual imagery is more analytically relevant. It is possible that crossmodal associations of the sort explored deeply in the work of Zohar Eitan and colleagues, and in the work of Steve Larson, might engender systematic relationships between acoustic characteristics and visual imagery such that analyses of musical structures could have something to say about affective response via the medium of visual imagery.\(^7\) But the kinds of visual imagery people experience in relation to music are highly variable with lots of individual differences based on extramusical factors, making this proposed mechanism a less likely candidate for identifying a tight coupling between musical structure and affective response.

Emotional contagion is arguably more clearly dependent on features of the music itself. Speech changes in systematic ways depending on the speaker’s emotional state,\(^8\) and music that mimicked these prosodic attributes (slow tempo, low pitch for sad, for example) could possibly elicit the corresponding emotional state by triggering auditory imagery. But if theorists care about affective response, this proposed mechanism poses a challenge to their ordinary modes of analysis, since the features that mimic prosodic expressiveness are almost exclusively those to which analysts attend the least (timbre, tempo, dynamics, etc.)—components usually relegated to the status of “secondary parameters.”\(^9\) Are the structural relationships analysts typically explore irrelevant to emotional experiences of music?

\(^6\) Juslin/Västfjäll 2008.  
\(^8\) Bachorowski 1999; Berckmoes/Vingerhoets 2004; Wennerstrom 2001.  
\(^9\) Meyer 1989, 14ff.
There is one last mechanism proposed by Juslin and Västfjäll that leaves room for significant connections between structure and affect: expectancy. If listeners can track patterns in music and forecast their likely continuations, and if violations and capitulations to these expectations can trigger affect, then patterns of the sort that interest analysts might be precisely those underlying the affective experiences of even everyday listeners. I will consider each of these conditionals in turn.

First, can listeners track patterns and forecast continuations? In answering this question, it is critical to be specific about what such tracking and forecasting might entail. Most listeners (theorists aside, perhaps) do not engage with music by explicitly attending to patterns and their likely continuations. If this kind of explicit tracking were required, we would have to answer no to this question and cross off expectancy as a viable mechanism for a structure-affect connection. But the past two decades have seen a stunning accumulation of knowledge about implicit processes in music perception—learning processes that occur outside of conscious awareness, such that listeners may disavow any relevant knowledge until experiments reveal it. In a classic example, well explored by Carol Krumhansl, most listeners will claim ignorance about tonal relationships, but will systematically rate the goodness of fit of probe tones to a context according to their position within the tonal system. Other classic findings include those of Jenny R. Saffran et al., which demonstrate that both infant and adult listeners track the statistical dependencies between tones in otherwise undifferentiated sequences with great accuracy but with absolutely no explicit awareness that they are doing so. The work of Mari Riess Jones has demonstrated that listeners allocate their attention selectively in time reflecting an implicit expectation for events to take place at specific future timepoints (on the beat, for example)—an idea that has been explored theoretically in the work of Christopher Hasty. Moreover, numerous studies using implicit measures such as reaction time and event-related potentials have demonstrated that people form expectations about likely continuations while listening to music.

So we know that listeners track patterns and forecast continuations, even though they may feel like they are thinking about dinner or about some sad event that occurred the last time they heard the piece. What about the second conditional—can violations and realizations of these expected continuations trigger affect? This idea’s origins lie in Meyer’s adaptation of John Dewey’s conflict theory of emotion, according to which the inhibition of tendency triggers affect. This theory has been modified and developed by a large number of people. But it seems undeniable that the expectancy-affect connec-

10 Reviewed in Ettlinger/Margulis/Wong 2011 and Rohrmeier/Rebuschat 2012.
11 Krumhansl 1990.
12 Saffran 1999.
14 Hasty 1997.
18 Reviewed in Huron/Margulis 2010.
tion, although the primary motivator for studying expectation in the first place, has been less satisfyingly characterized than the expectations themselves.

**Surprise and Dimensionality**

One challenge of expectancy-related music research has been that unidimensional characterizations of expectation have led to unidimensional characterizations of affect, with musical experience depicted as a series of more and less intense surprises. In recent years, more sophisticated connections have been drawn between expectancy, expectancy violation, and the very real senses of tension and relaxation that can partly define the moment-to-moment experience of music. But what is left over after tension and relaxation are accounted for is a lot—for one thing, affective experience can be highly differentiated. Music can seem sad or exuberant, resigned or vainglorious. Must these percepts be left for other mechanisms to explain or might expectancy contribute to shaping these differentiated kinds of responses as well?

In a study published as a chapter in the *Handbook of Topic Theory*, I pursue a hypothesis relating to the role of context in the differentiation of surprise-based affect. This study looks to work from another corner of music theory, namely topic theory, to understand the ways context might predispose listeners to interpret surprise in different lights. The same syntactic surprise—a general pause—was inserted after a cadential 6/4 in excerpts featuring one of four different topics with distinct affective connotations. Participants heard, in randomized order, two excerpts from each of these four topical categories in two conditions: expected (without a general pause after the cadential 6/4) and surprising (with a general pause after the cadential 6/4). They heard each of these excerpts four times across the course of the session. During each hearing, they continuously rated the piece along a single affective dimension; for example, on one hearing they rated how playful the piece seemed at each moment, but on another they rated how ominous it seemed at each moment. By comparing ratings at the moment of surprise (the general pause) with ratings at the corresponding moment in the expected version, the contribution of the surprising event to perceptions along that dimension could be assessed. Results showed that the same syntactic surprise (the general pause) triggered different affective interpretations in different topical contexts. For example, in a context featuring the *brilliant style*, the surprise might selectively elevate impressions of playfulness, but in a context featuring the topic *siciliano*, the surprise might selectively elevate impressions of ominousness.

The implication of this finding is that context might contribute to differentiating surprise-based affect. Surprise may intensify affective response, but whether the intensified impression is one of playfulness or ominousness or anything else depends on the appropriate contextual priming. Although the study described above examines context in the narrow sense of eighteenth-century musical topics, the same principle could ostensibly be at work in any repertoire where conventional association, emotional contagion, or

---

20 Margulis forthcoming.
some other mechanism\textsuperscript{21} connects musical structure to particular affective realms. My study argues that stylistic context can set the appropriate affective territory, with expectation contributing dynamically to the modulation of expressive intensity within this established dimension.\textsuperscript{22}

This account leaves room for individual differences in affective response. As I argued in one of my previous articles, any theory that attempts to relate expectation and listening experience needs to be specific and explicit about

(1) the expectation’s origin—where does the expectation come from and why might a listener have it in the first place;
(2) the expectation’s nature—what kind of expectation is it and what does it feel like to have it;
(3) the expectation’s time course—does the expectation target a specific event at a specific time or relate to some more temporally extended characteristic of a passage;
(4) the expectation’s object—what kind of thing does the expectation predict;
(5) the expectation’s consequence—what effect does the expectation have on things we might care about as analysts, listeners, or psychologists.\textsuperscript{23}

If the consequence of the expectation is a fleeting perception of intensification along the affective dimension established by the preceding context, then something idiosyncratic about the way a particular listener interprets a passage could transform the phenomenology of the surprise. For example, if a listener had heard processional music most frequently at graduation ceremonies, music in this style might trigger associations with nostalgia and wistful sentimentality, but if a listener had heard processional music mostly at funerals, music in this style might trigger associations with loss and grief, via the mechanism of evaluative conditioning outlined by Juslin and Västfjäll.\textsuperscript{24} A syntactically surprising event might be experienced by the first listener as a moment of extra wistfulness, but by the second listener as a moment of intensified sadness. In common between these two hypothetical listeners would be an experience of intensity at the moment of surprise, but different between them would be the specific phenomenology of this intensity.

With this account we come full circle back to Meyer’s assertion that expectation allows the analyst to factor out the listener and look directly at the relationship between musical structure and affect. Close scrutiny of expectation may indeed allow the analyst to engage with affective response, but only by admitting the listener into the relationship. Wong et al., for example, show that listeners enculturated in Western music experience melodies on the sitar as tenser than the same melodies played on the piano, but listeners enculturated in Indian music experience melodies on the piano as tenser than the same melodies played on the sitar.\textsuperscript{25} Experience and enculturation shape experience in im-

\textsuperscript{21} See Juslin/Västfjäll 2008.
\textsuperscript{22} Margulis forthcoming.
\textsuperscript{23} Margulis 2007, 205f.
\textsuperscript{24} See Juslin/Västfjäll 2008, 564f.
important ways. I propose that Meyer’s hypothesized structure-affect link and the affective variability that results from enculturation and other factors can cohabitate in the following way: the moment-to-moment dynamics of expectation-based affective response are broadly shared by listeners, but contextual affective associations vary more from person to person. Thus it might be possible to talk about musical expectations without having to accommodate the affective response of each individual listener, while still integrating specific knowledge about particular listeners’ background experience in other ways.26

This sort of expectational theory aims to identify the mechanisms and processes that underlie reactions to music that a person might sustain regardless of whether or not they read an analytic account. A listener might experience a twinge of foreboding at a particular moment, for example, and the theory might be able to identify some link between the context and this affective terrain as well as a syntactic surprise at the relevant point, accounting together for the character and the timing of this perception. In other words, the aim of the analysis is to explain some response that a listener might have, rather than to suggest some new response to the listener—it is descriptive rather than prescriptive.

For example, the second movement of Haydn’s Symphony No. 64 features a general pause after the cadential 6/4 in m. 4. According to the theory outlined here, this syntactic surprise elevates intensity along some affective dimension—it affects the dynamics of the affective response. The content of the affective response—the dimension along which it is experienced—might vary according to the style of the surrounding musical context and the particular listening background of the individual listener. In my study, the opening of this movement was characterized as representative of the singing style, and listeners tended to register an elevation in impressions of sublimity across the course of the pause. But when this same pause was inserted after the cadential 6/4 in other excerpts, listeners interpreted it as intensifying very different affective impressions, such as ominousness or playfulness. The pause always elevated the intensity of some affective impression, but which affective dimension could vary.

Ineffability and the Explanatory Force of Analysis

An important strand in theoretical, philosophical, and historical perspectives on music has emphasized some incontrovertibly ineffable aspect to musical experience. Diana Raffman focused on the importance of expressive nuance, and Mark DeBellis focused on the resistance of many aspects of listening to conceptual capture.27 An English translation brought Vladimir Jankélévitch’s thought on this subject to a broader audience, as did an associated article by Carolyn Abbate in which she emphasized the essentially “drastic,” sensory nature of musical experience over its various “gnostic” or abstracted interpretations.28 Indeed, there is something about actually listening to music, rather than abstract-

26 Margulis 2014b.
ing or thinking about it, that compels us to visit and revisit favorite pieces.29 Whereas linguistic narratives tend to elicit gist rather than verbatim memory30—memory devoted to the events recounted by the narrative rather than the specific words used to encode it—this seems not to be the case for music. Once I know what happens in a particular story, I might not be interested in reading it again and again. What is more, depending on the type of narrative, I might be able to hear a summary and skip the whole story altogether, satisfied that I would have absorbed the main points. But it is almost impossible to imagine a scenario under which a summary of a piece of music would be accepted as a sufficient proxy for listening to it. There is something in the dynamic, moment-to-moment experience of listening to music that defies conceptualization and summary.

This defiance, in my view, provides the very *raison d'être* for music analysis. In addition, it positions cognitive science at the center rather than the periphery of the analytic enterprise. One thing analysis tries to do is find a vocabulary to talk about experiences that are inherently resistant to articulation. This is what many of analysis’ most familiar terms are trying to do: “prolongation,” for example, attempts to describe a subtle experience wherein a particular musical event continues to exert imagined influence even as other events succeed it in time31; “downbeat” refers to an equally subtle experience where a particular timepoint elicits more attention or emphasis than surrounding ones, and causes these surrounding timepoints to be understood relationally to it.32 Another thing analysis tries to do is identify the mechanisms that might give rise to these experiences. Since these mechanisms are mostly nontransparent—that is, we lack explicit access to them, and remain privy only to their effects—we need methodologies that will allow us to peer inside the black box of the mind. Cognitive science, as a discipline, is devoted to just such peering, and has developed a host of sophisticated tools for opening up the mind’s black box and exposing its inner workings.

According to this view, introspection is not something hopelessly solipsistic in which analysts occasionally indulge; rather, it is a fundamental part of a process that might also include some empirical or quantitative component. If we are committed to the notion that important aspects of musical experience are hard to talk about, then before we can explain these aspects, or understand the mechanisms underlying them, we need to make these elusive perceptual experiences available for discussion. That involves a committed kind of introspection.

Consider, for example, a passage from Lawrence Kramer’s interpretation of the opening of an inconspicuous piece from Robert Schumann’s *Album für die Jugend*, No. 34—*Thema*. Kramer talks specifically about the way the augmented chords in this piece have “a particular air of dissolving the harmony as they enhance the disorienting effect of the diminished chords,”33 and more generally about the way that the theme’s tonality is po-

---

29 This proclivity is explored in Margulis 2013 and Margulis 2014a.
30 Reyna/Brainerd 1995.
31 Larson 1997 beautifully explores the perceptual attributes of this phenomenon.
32 This experience is sensitively described by Hasty 1997, and Epstein 1995, among others.
33 Kramer 1981, 194.
sitioned as a horizon rather than an overt presentation. These may seem like subjective impressions, opposed by nature to the kind of conclusions producible by empirical study.

But when I read Kramer’s lines, I experience a sense of recognition, as if Kramer had wrested into articulable form impressions I was already having when listening to the piece. The sense of recognition that emerges when reading satisfying music analysis implies that even very subtle aspects of the musical experience are broadly shared across listeners, and this commonality implies that there is something systematic going on that generates these impressions. Kramer’s analysis identifies impressions of dissolution, disorientation, and obliqueness (in the sense that tonality is referred to rather than straightforwardly presented). These characterizations are successful if the analysis’s reader experiences a sense of recognition. It is as if a patient with a hard-to-describe headache saw a doctor who asked in turn if the pain was sharp and stabbing or dull and constant, and when the doctor hit upon the headache’s characteristics, the patient responded with excitement “yes, that’s exactly how it feels!” The doctor could then say, for example, “aha—that kind of pain sounds like migraine, which is caused by constriction of blood vessels leading into the brain.” The doctor has performed two services here: she has given the patient a way of describing an experience he would already been having, and she has explained the mechanism by which this experience is generated.

Analysts often perform this critical first service—they provide a way of talking and thinking about something that is challenging to talk and think about it. Once an experience has been delineated or made available to thought and discussion in this way, it becomes possible to explore the mechanisms that might underlie it. It is at this stage that cognitive science can offer a wealth of relevant perspectives and methodologies. What causes impressions of disorientation and obliqueness? There are parallels in social interaction, where an absence of entrainment, a rhythmic out-of-phasedness can trigger sensations of disorientation, and to the obliqueness that characterizes conversational exchanges where the real topic at hand is only danced around, and never explicitly referred to. Linguists and psychologists know a lot about these kinds of situations and the mechanisms that make them possible. Adapting these experimental paradigms to musical stimuli could result in insight into the ways that even subtle, nuanced, dynamic aspects of the listening experience arise.

It is possible that the impressions Kramer identifies would only arise in the minds of “experienced listeners,” in the sense of the term described at the start of this article. Listeners unfamiliar with common practice idioms and conventions may experience something altogether different. Within the broad swath of listeners who would qualify as experienced, subsets may experience different kinds of associated visual imagery or valenced impressions depending on prior varieties of evaluative conditioning. Neither cognitive science nor analysis will ever explain every part of any particular listener’s experience; however, in partnership they may not only identify interesting areas of overlap but also provide insight into the mechanisms that give rise to these shared experiences. Expectation is a particularly promising proposed mechanism, with a host of associated methodologies and a clear relevance to the way music unfolds in time. As music theory and cognitive science develop closer relationships in the future, it is likely that the number of mechanisms understood to link musical structure and musical affect extend well beyond those currently proposed.
References


