Music Theory, Multimedia, and the Construction of Meaning

Lawrence M. Zbikowski

In the summer of 1938, as the storm clouds of war were gathering across Europe, Sir Donald Tovey delivered a lecture to the British Academy entitled "The Main Stream of Music." The lecture is a curious affair, not the least because for Tovey the mainstream of music was a thoroughly Germanic one. While sensitive to the accomplishments of non-German composers in the sixteenth and seventeenth centuries, Tovey nonetheless believed there was a sea change in musical composition in the early eighteenth century: "With the advent of Bach, music became an art so congenial to all that is best in the Teutonic intellect that for the next two centuries there is no musical art-form in which German musicians have not produced the supreme masterpieces."[1] And it was the genius of Bach that discovered resources within music which rendered the medium independent of other media. Tovey continues, "There can be no supreme musical art without the qualities of absolute music, whether the art be as compounded with other arts as Wagnerian opera or as exclusively musical as the string quartets of Beethoven." The mainstream of music, then, was one flooded by the works of German composers, works whose excellence relied on the purely musical.

This conclusion caused Tovey some anxiety. Indeed, both his long-held belief that music could speak to a broad audience, and his tireless championing of British music, were challenged by a central corpus of thoroughly German works that required neither text nor program for their understanding. But a deeper source of his anxiety was a nagging suspicion that musicians were in danger of losing their way. Some pages later, after having drawn his survey to a close with a brief contemplation of Wagner's enormous operas,

[1] Tovey 1938: 128.

he writes "I can go no further. At the present day all musicians feel more or less at sea, and not all of us are good sailors."[2]

Sixty-five years later one can only look with envy on the navigation problem that confronted Tovey, for his mainstream is now regarded by most as but a tributary, if a significant one, to the vast body of music through which scholars must find their way. This challenge to navigation is, in less metaphorical terms, a challenge to musical analysis, for analysis is one of the fundamental ways musicians chart their course through challenging or unfamiliar repertoire. And one seldom finds a repertoire that presents as many challenging or unfamiliar problems as does musical multimedia, for the various ways music can combine with words or images yield phenomena that are often beyond the reach of our usual analytical tools. Indeed, as Nicholas Cook argues in *Analysing Musical Multimedia*, confronting multimedia opens up basic issues within the theory and analysis of music, and suggests a thorough re-evaluation of the entire enterprise. As Cook notes, "What begins as an analysis of musical multimedia, then, turns ineluctably into an analysis of analysis" (viii).

The analysis of analysis begins not with the somewhat shopworn questions of what counts as analysis and why one should do it, but with the issue of musical meaning, for the assumption that music means *something* is basic to musical multimedia. This is not to say that musical meaning is theorized in any profound way by those who create musical multimedia, only that these practitioners realize that a television commercial or a film means something quite different when the music is taken away or substantially altered. Thus, while music often occupies a place well below the obvious story-line within these media, its contribution is not inconsequential—as Cook observes, "Music transfers its own attributes to the story-line and to the product; it creates coherence, making connections that are not there in the words or pictures; it even engenders meanings of its own" (20). This leads Cook to the somewhat startling conclusion that music in the abstract—Tovey's "absolute music"—doesn't have meaning.

---

[2] Tovey 1938: 139.

> What it has, rather, is a *potential* for the construction or negotiation of meaning in specific contexts. It is a bundle of generic attributes in search of an object. Or it might be described as a structured semantic space, a privileged site for the negotiation of meaning. And if, in the commercials, meaning emerges from the mutual interaction of music, words, and pictures, then, at the same time, it is meaning that forms the common currency among these elements—that makes the negotiation possible, so to speak. (23)

Cook goes on to argue that the same holds true for the words and music in songs, and the words *about* music in analytical prose—in all cases, the meaning that is produced is a consequence of interactions between various media. Musical culture is, in consequence, irreducibly multimedia in nature (23). Analysis must perforce deal not only with the interaction between musical elements but also with the interactions between media, for these interactions are basic to the construction of meaning.

The interactions between media that Cook sees as most important are oppositional in nature—what is significant is not how media are *like* one another, but how they are *different* from one another. This sense of discrete media that in some way interact is, Cook argues, what separates the experience of multimedia from synaesthesia. At the same time, the most compelling examples of multimedia are not simply the consequence of the coincidence of two discrete forms of communication. What is required is a limited intersection of attributes between the constituent media—what Cook calls an enabling similarity—which allows the media to be brought together into the same conceptual domain so that their differences can be noted and thus made accessible for the construction of meaning.[3]

This notion of domains that are in some respects similar which are brought into a correlation that reveals their differences brings Cook to the theory of metaphor first proposed by George Lakoff and Mark Johnson in the early 1980s.[4] In the following, I would like to explore the contemporary theory of metaphor in just a bit more detail than Cook is able to do in *Analysing Musical Multimedia*. Further developments of this theory offer ways to

[3] A similar perspective, developed from research in psychology, can be seen in Gentner and Markman 1994 and 1997.

[4] Lakoff and Johnson 1980.

streamline a few aspects of Cook's account of multimedia, and extensions to the theory offer a somewhat more systematic approach to the analysis of multimedia in particular, and music in general.

The Contemporary Theory of Metaphor

Lakoff and Johnson's point of departure was the proposal that metaphor was not simply a manifestation of the figural use of language to create colorful if imprecise images but reflected a basic structure of human understanding.[5] For instance, in speaking about a person's romantic relationships we might use expressions such as "He is known for his many rapid conquests" or "She is besieged by suitors." The linguistic metaphors central to these expressions are based on the conceptual metaphor LOVE IS WAR, which correlates the conceptual domain of romantic love with the conceptual domain of warfare. Once this correlation is active we can access concepts drawn from the domain of warfare ("rapid conquests," "besieged") to characterize aspects of individuals' romantic relationships. More generally, WAR serves as a source domain, providing a rich set of structures that we can map onto the target domain of LOVE. Thus "quickly bringing an enemy to defeat" is used to structure our understanding of a situation in which an individual is able to cause other individuals to direct their affections only to him, and to do so with little effort: "He is known for his many rapid conquests."

One question raised by this approach to metaphor was of the ultimate grounding of the process of mapping structure from one domain onto another. Even if we grant that we understand a target domain (such as LOVE) in terms of a source domain (such as WAR), how is it that we understand the source domain in the first place? Mark Johnson answered this question by proposing that meaning was grounded in repeated patterns of bodily experience, which give

[5] Expanded versions of the discussion that follows, along with more extensive citations to recent work on metaphor theory, can be found in Zbikowski 1998 and Zbikowski 2002: 65-71.

rise to what he called image schemata.[6] An image schema is a dynamic cognitive construct that functions somewhat like the abstract structure of an image and thereby connects together a vast range of different experiences that manifest this same recurring structure. Thus our understanding of a source domain like WAR is grounded in image schemata such as BLOCKAGE and COUNTERFORCE; these, together with evaluative judgments such as "winning" and "losing," provide a rich conceptual structure which can then be mapped onto domains such as LOVE.

Although the theory of image schemata provides a way to explain how cross-domain mapping is grounded, it does not explain why some mappings are more felicitous than others. For instance, we could map structure from the domain of WAR onto the domain of PITCH RELATIONSHIPS to produce statements like "The G4 vanquished the F♯4." But if we simply want to describe how one pitch relates to another this seems a bit much—we tend to prefer mapping from the domain of ORIENTATION IN PHYSICAL SPACE: "The G4 is higher than the F♯4." To account for why some metaphorical mappings are more effective than others, George Lakoff and Mark Turner proposed that such mappings are not about the *imposition* of the structure of the source domain on the target domain, but are instead about the establishment of correspondences between the two domains. These correspondences are not haphazard, but instead preserve the image-schematic structure latent in each domain. Lakoff and Turner formalized this perspective with the Invariance Principle, which Turner states as follows: "In metaphoric mapping, for those components of the source and target domains determined to be involved in the mapping, preserve the image-schematic structure of the target, and import as much image-schematic structure from the source as is consistent with that preservation."[7] Our mapping of orientation in physical space onto pitch thus relies on correspondences between the image-schematic structure of components of the spatial and acoustical domains. Both space and the frequency spectrum are continua that can be divided into discontinuous elements. In the spatial domain, division of the continuum results in points; in the

[6] Johnson 1987.

[7] Turner 1990: 254; emphasized in original. See also Lakoff 1990.

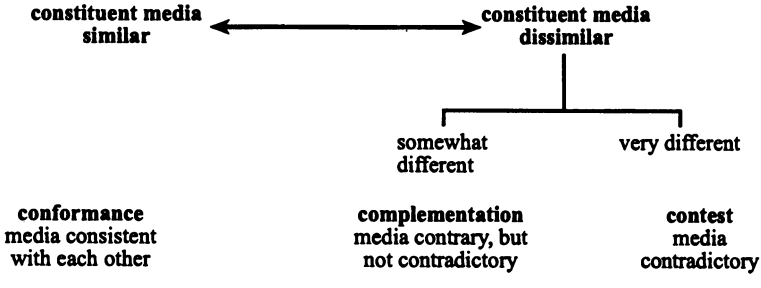
acoustic domain, it results in pitches. The mapping thus allows us to import the concrete relationships through which we understand physical space into the domain of music and thereby provide a coherent account of relationships between musical pitches. In contrast, mapping from the domain of WAR onto the domain of PITCH RELATIONSHIPS works less well because it does not preserve the image-schematic structure of the target domain (our sense that the frequency spectrum is a continuum is almost completely suppressed) and because it imports structure (notions based on BLOCKAGE and COUNTERFORCE) foreign to the target domain.[8]

According to the contemporary theory of metaphor, then, metaphor is a basic cognitive capacity that involves mapping structure from one domain onto another. This mapping is possible because there are aspects of the structure of each domain that are invariant—these are the enabling similarities that Cook suggests are a precondition for musical multimedia. Thus, in the case of Schoenberg's *Die glückliche Hand* (discussed by Cook on pp. 41-56), the "Lighting Crescendo" that occurs in bars 125-53 relies on shared structure between the music, lighting, and action on the stage. As the musical materials get louder and coalesce the lighting gradually goes from dull red through a variety of hues until it becomes a glaring yellow, and the central character moves from a portrayal of exhaustion through stages that lead to a portrayal of extreme tension. The basic structure that unites these three domains relies on the notion of gradually increasing energy. The instantiation of this structure in each domain makes it possible for the media to combine; because the structure is instantiated differently within each domain the result of the combination is *multi*media. An increase in energy such as that portrayed by the actor might well be soundless, but here it is accompanied by a crescendo and the emergence of musical themes from an inchoate background; that same increase in energy might well play out within consistent lighting, but in Schoenberg's conception it begins in murky gloom and ends in the bright light of day.

[8] Note, however, that if our concern were tonal relationships as opposed to pitch relationships a mapping from the domain of WAR might be completely appropriate. See Burnham 1995, Chap.1.

The perspective provided by metaphor theory leads Cook, at the conclusion of the first part of his book, to propose three basic models for multimedia. The models are shown on Example 1, which places them along a continuum which focuses on the relative degree of similarity among the constituent media of an instance of multimedia, or IMM. Leftmost on the diagram is the conformance model, distinguished by the large number of similarities that obtain between the constituent media of an IMM. Differences between the media are thus relatively attenuated, and in extreme cases a conformance IMM might be taken as an instance of a single medium. On the right of the diagram is the contest model, in which similarity obtains at only the most abstract level. Differences between the media are thus profound, and in extreme cases a contest IMM will simply break apart into its constituent media. These media would still be coincident, but they would not yield an IMM. In between these two extremes is the complementation model: differences between the constituent media of the IMM are significant enough that the media can be readily distinguished from one another, but not so marked that the media seem to contradict each other.

Example 1. Cook's three models of musical multimedia situated along a similarity-dissimilarity continuum.



Cook is thus able to extract a promising analytical approach to musical multimedia from contemporary theories of metaphor. This approach allows him to characterize the conditions that will yield an instance of multimedia, and to develop a typology of such instances based on similarities between their constituent media. However, the approach also raises two problems, both of which stem from limitations of the contemporary theory of metaphor. First, the emphasis in metaphor theory has been on mapping structure from one domain onto another. While describing what music contributes to our understanding of the stage action in Schoenberg's *Die glückliche Hand* is an important first step in understanding that particular IMM (since it allows us to specify what structures from the musical domain are mapped onto the domain of the stage action), it does not yield a description of the IMM itself. That is because we also want to know which structures from the domain of the stage action are mapped onto the musical domain. The same holds true for mappings between the domains set up by the music and the lighting, and between the domains set up by the stage action and the lighting.

Second (and related), metaphor theory offers no account of the unique conceptual domain that some cross-domain mappings produce. While mappings between the domains of music, stage action, and lighting are important for the process of meaning construction initiated by bars 125-53 of *Die glückliche Hand*, the unique domain that these mappings produce—the instance of multimedia specific to this moment in Schoenberg's monodrama—is what we are really interested in. But it is just this specification of elements and relations proper to the IMM that is lacking in accounts that focus only mapping structure from a source domain onto a target domain.

These two limitations of the contemporary theory of metaphor—the difficulty of accounting for coordinate mappings between two or more domains, and acknowledgement that such mappings often yield new conceptual domains—gave rise to the theory of conceptual blending. In the following I outline basic features of this theory, and describe its application to the analysis of musical multimedia.

Conceptual Blending and Musical Multimedia

A conceptual blend begins with concepts drawn from two correlated domains. Consider, for instance, Marcel Proust's recollection of one feature of the springtime walks along the "Méséglise way" during visits to Combray:

> We would leave town by the road which ran along the white fence of M. Swann's park. Before reaching it we would be met on our way by the scent of his lilaç-trees, come out to welcome strangers. From amid the fresh little green hearts of their foliage they raised inquisitively over the fence of the park their plumes of white or mauve blossoms, which glowed, even in the shade, with the sunlight in which they had bathed.[9]

Proust's description relies on concepts drawn from the domain of trees (including not only concepts associated with the scent, foliage, and blossoms of trees, but also with their shape and disposition) and from the domain of intelligent beings (including concepts associated with welcoming strangers, being inquisitive, and bathing). These concepts are then blended together to create a domain in which the lilac trees are more than alive—they are also intelligent and animate. Within this domain there are new structures that cannot be found in either of the two original domains. In the blended space, the lilacs send forth their scent, raise their foliage inquisitively over the fence, and bathe in sunlight.

In order to study conceptual blends such as that represented by Proust's description, the rhetorician Mark Turner and the linguist Gilles Fauconnier developed the notion of conceptual integration networks (CINs).[10] Each CIN consists of at least four circumscribed and transitory domains called mental spaces. Mental spaces temporarily recruit structure from more-generic conceptual domains in response to immediate circumstances and are constantly

---

[9] Proust 1981: 147-148.

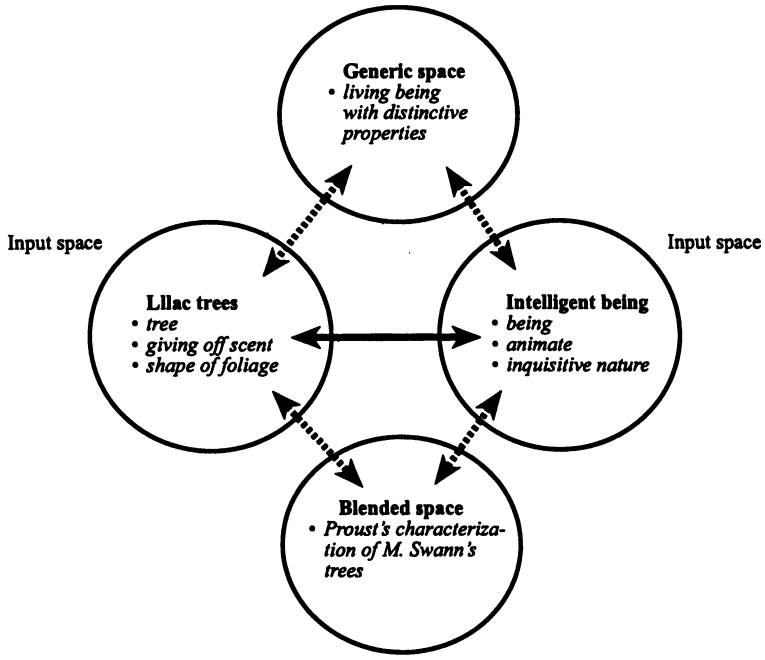
[10] I provide an overview of work on conceptual blending, and its application to music, in Zbikowski 2002: 77-95. The most comprehensive study of conceptual blending as of this writing is Fauconnier and Turner 2002. Cook has also made use of blending theory; see Cook 2001.

modified as our thought unfolds.[11] For instance, Proust's description of the walk along M. Swann's park sets up two correlated mental spaces. The first is that of the lilac trees, the second that of intelligent beings. Features of these two spaces are combined in a third mental space, producing the intelligent and animate trees of Proust's description. Turner and Fauconnier use CINs to formalize the relationships between the mental spaces involved in a conceptual blend, to specify what aspects of the input spaces are imported into the blend, and to describe the emergent structure that results from the process of conceptual blending.

The CIN for the conceptual blend summoned by Proust is diagrammed in Example 2. The network involves four interconnected mental spaces, which are shown as circles. Central to the network are two correlated input spaces, the "lilac trees" space and the "intelligent being" space. The solid double-headed arrow linking these two spaces indicates that elements within them serve as structural correlates: *tree* is correlated with *being, giving off scent* with *animate,* and *shape of foliage* with *inquisitive nature.* Guiding the process of mapping between these spaces is the generic space, which maps onto each of the input spaces and contains what they have in common: a living being that we come to know on account of certain important properties. Guided by the conceptual framework provided by the generic space, structure from each of the input spaces is projected into the fourth space, called the blend, which yields Proust's characterization of M. Swann's lilac trees. The mapping is only partial, however, reflecting the limitations imposed by the generic space. Since the generic space is not concerned with incidental properties, such as the means by which living beings take nourishment, the sunlight essential for the trees' continued life is relegated to the role of a simple sensual pleasure within the blend—that is, Proust does not liken the sun in which the trees have bathed to a good meal.

[11] The theory of mental spaces is developed in Fauconnier 1994; Fauconnier 1997; Fauconnier and Sweetser 1996; see also Turner 1996.

Example 2. CIN for Proust's characterization of M. Swann's lilac trees.



The dashed arrows linking the generic space to the input spaces, and the input spaces to the blended space, indicate the directions in which structure is projected: from the generic space to the input spaces, and from the input spaces to the blended space. The arrows are double-headed because, under certain circumstances, structure may also be projected from the blended space back into the input spaces, and from the input spaces back into the generic space. The double-headed arrows also serve as a reminder of the limitation of all of the diagrams of CINs I shall use: mental spaces are dynamic structures, as are the CINs that are built from them. What Example 2 represents is a sort of analytical snapshot of this particular network, framed with the intent of capturing its essential features, but making no claim to exhausting the possibilities for description. Hints about how the CIN and its

spaces may develop can be gleaned from the diagram, but a full account would require a series of such snapshots.

Two important features of the process of conceptual blending are illustrated by this example. First, new structure emerges in the blend. For instance, the inquisitive nature typical of intelligent beings combines with the shape of the lilac trees' foliage to yield trees that reach out toward the visitor. Given the various combinations of concepts we can easily complete the picture Proust has sketched for us, and imagine engaging in a dialogue with the trees. And we can also use the blended concepts to elaborate the story, and imagine that the lilac trees are but one of a number of intelligent plants that populate M. Swann's park. All of this structure is specific to the blended space, made possible once we have correlated the two input spaces.[12]

Second, blending allows Proust to tell a complex, multidimensional story—one that extends to the sights, sounds, and lived experience of springtime—in a highly compressed version that focuses on a single human-like form.[13] Proust's blend takes a season that stretches over weeks if not months, that is manifested in a tremendous outburst of renewal and change, and compresses it into a single encounter with an imaginary individual.

Evidence suggests that conceptual blending is pervasive in human understanding, and that blends can be much more complex than the one exemplified by Proust's vignette. There is, accordingly, much more to the theory of conceptual blending than what I have sketched here. Nonetheless, only two further aspects of this theory need concern us here: conceptual blends with multiple input spaces, and the technical resources offered by blending theory.

Although introductions to conceptual blending often concentrate on the four-space model illustrated in Example 2, conceptual blends typically involve more than two input spaces, yielding CINs with five or more spaces. The situation is well

---

[12] Further discussion of the three operations important to conceptual blending that I have outlined here—composition, completion, and elaboration—can be found in Fauconnier and Turner 2002: 48-49 and Zbikowski 2002: 77-95.

[13] "Story" as I use it here refers to the rich, embodied parabolic structures that Mark Turner has argued are essential to human thought; see Turner 1996.

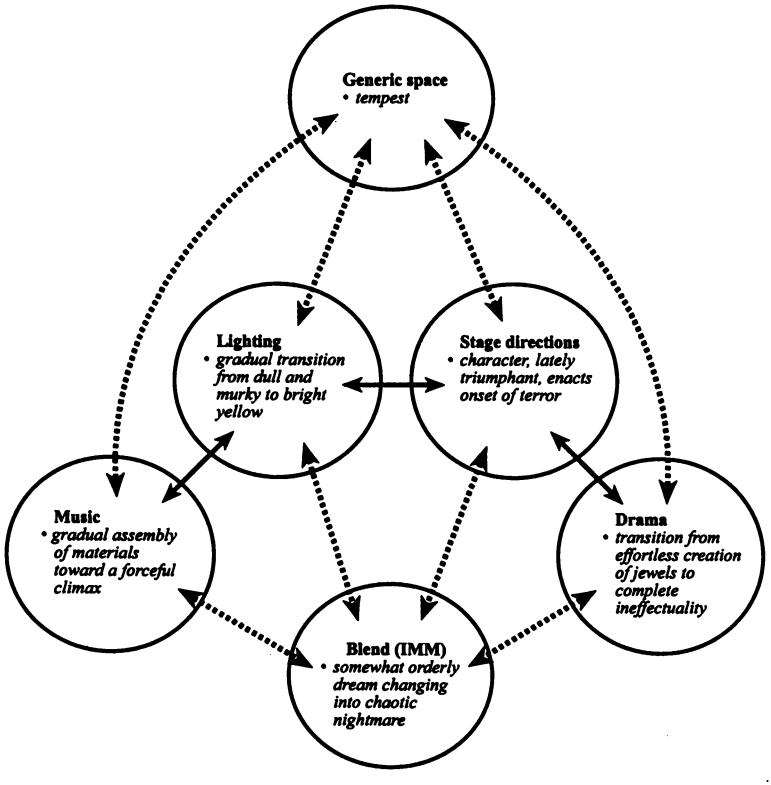
illustrated by the lighting crescendo from *Die glückliche Hand.* Here at least four input spaces are involved: one is set up by the music, another by the lighting, a third by the actions that take place on stage, and the final space by the dramatic story told by the text that is sung.[14] The blend that results is, generically, an instance of multimedia, but more specifically it is a compelling portrayal of the moment when a dream turns into a nightmare.

The essential features of the spaces proper to the blend are diagrammed in Example 3. The music space is occupied with a gradual assemblage of diverse musical fragments, sounding across a large orchestra, which eventually coalesce into a significant and inexorable force. The lighting space is built around the gradual transition from murky gloom to bright daylight noted above. The stage directions set up a space in which the central character, who moments before had created finished jewels with a single blow of a hammer, now finds himself in the thrall of incomprehensible and terrible powers. And within the dramatic space set up by the text there is a sense of anticipation: the man's "Das kann man einfacher!" ("That can be done more simply!," spoken upon seeing workers laboriously making jewels) and "So schafft man Schmuck!" ("This is the way to make jewels!," spoken after the hammer blow) have taken him far from the misery he was in at the opening of the drama. Nonetheless, his inability to completely throw off that misery—his clothes are still ragged, and he still seems to wander rootlessly—suggests that this distance cannot be maintained.

Elements from each of the four spaces are projected into the blended space, guided by the central image of a tempest (which is how Schoenberg refers to the cataclysm). From the mental space set up by the music are projected the gradual assembly of disparate sonic elements and a dynamic sweep that comprises both an increase in timbral resources and an increase in volume. There is a

[14] Above I discuss only three correlated domains within *Die glückliche Hand,* in line with Cook's discussion on pp. 41-56; here I have added a fourth. And Schoenberg at least suggests that there might be additional spaces: he calls for a wind machine at m. 129 (that is, a sound generating device separate from the instruments), but he specifies that the sound of the wind should not cover the sound of the orchestra. It is, however, a bit unclear whether the sound effect would prompt the construction of yet a further space (for quasi-natural sounds) or whether it would be absorbed into the overall sounds of the orchestra.

Example 3. CIN for bars 125-53 of Schoenberg's Die glückliche Hand.



similar dynamic sweep projected from the space set up by the lighting, but it is the inverse of what we might usually expect: the tempest begins in murky darkness and climaxes in the bright light of day. The increase in intensity is similar to what happens as a tempest gathers power, but the final result—glaring yellow light—is not strongly correlated with our impressions of storms. Rather, visual impressions of a body reacting to a tempest are projected into the blend from the space set up by the stage directions, as is the sense of anticipation from the dramatic space. The instance of multimedia that results from combining these various projections within the blended space is reminiscent of a tempest, but not one that comes from without: although the noise and tumult of the music and the sense of dramatic anticipation that derives from the stage directions and dramatic situation would support this, the receding darkness and restrained (if nonetheless tense) gestures by the actor speak of internal tumult. This is a tempest that arises from within, an impression that conforms with Schoenberg's stage directions: "Der Mann hat dieses Crescendo des Lichts und des Sturmes so darzustellen, als ginge beides von ihm aus" ("During the crescendo of lights and storm the man reacts as though both emanated from him").[15]

The second aspect of blending relevant for the analysis of musical multimedia concerns patterns within CINs. Although networks can take diverse forms, and can—as the foregoing discussion illustrated—involve multiple input spaces, most blends can be characterized in terms of one of four types of CINs: simplex networks, mirror networks, single-scope networks, and double-scope networks.[16] Most IMMs can be described in terms of the latter three. For instance, among the three IMMs represented in Example 1, the conformance model can be characterized as a

[15] I should note that the analysis I offer here is intended to illustrate a conceptual blend that involves multiple input spaces. Cook discusses evidence that Schoenberg did not regard all of these "input domains" as equivalent, suggesting a differential projection from the various source domains into the blend. Such differential projection is well recognized within blending theory; see Fauconnier and Turner 2002.

[16] Each of the four main kinds of CINs are discussed in detail in Fauconnier and Turner 2002: 119-135.

mirror network: because the input spaces share a common organizing frame they "mirror" one another, and differences between the inputs are greatly attenuated. The contest model, on the other hand, is a classic double-scope network: the input spaces have incommensurate organizing frames, and differences between the inputs are consequently accentuated. Finally, the complementation model can be characterized as a single-scope network: because the organizing frame for one of the inputs typically dominates the IMM, differences between the inputs are somewhat muted but still noticeable. In general, the technology of conceptual blending—which gives ways to describe what materials are projected into the blend (or IMM) and what the relationships are between the various mental spaces within the network—makes it possible to give detailed characterizations of the similarities *and* differences between the constituent media of an IMM. Conceptual integration networks, then, offer a way to simplify and make consistent technical descriptions of relationships between the constituent media of an IMM, and to allow for comparisons between IMMs and other conceptual blends.

Conclusion

Although Tovey's survey of the mainstream of music left him feeling somewhat at sea, the waters he navigated were nonetheless familiar ones, and ran deep with absolute music. Today the notion of absolute music has at best heuristic value; at worst, it obscures the entire enterprise of the scholarly study of music, for it strips musical practice of the cultural and social context that makes music meaningful. This perspective compounds the problems modern scholars face in navigating the vast body of music before them, for success requires that they look beyond "the music itself." As Cook shows, analyzing musical multimedia is a kind of laboratory for studying these problems because it pushes the issue of musical meaning, and the role of analysis in revealing that meaning, to the foreground. If music indeed contributes meaning to a television commercial or a film, what is the nature of this meaning, how can we describe it, and how does it connect with the sequence of sound phenomena we recognize as properly musical? Although theories of

cross-domain mapping and conceptual blending cannot answer all the questions we would ask about musical meaning, they do suggest that these issues are not unique to musical understanding, and they provide methodologies for exploring the relationship between music and other media. We may, with Tovey, still be more or less at sea, but the firmament may not be quite as distant as we once thought.

References

Burnham, S. 1995. *Beethoven Hero.* Princeton, New Jersey: Princeton University Press.

Cook, N. 2001. "Theorizing Musical Meaning." *Music Theory Spectrum* 23/2 (Fall): 170–95.

Fauconnier, G. 1994. *Mental Spaces: Aspects of Meaning Construction in Natural Language.* 2d ed. Foreword by G. Lakoff and E. Sweetser. Cambridge: Cambridge University Press.

_____. 1997. *Mappings in Thought and Language.* Cambridge: Cambridge University Press.

Fauconnier, G., and E. Sweetser, eds. 1996. *Spaces, Worlds, and Grammar.* Chicago: University of Chicago Press.

Fauconnier, G., and M. Turner. 2002. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities.* New York: Basic Books.

Gentner, D., and A. B. Markman. 1994. "Structural Alignment in Comparison: No Difference Without Similarity." *Psychological Science* 5/3 (May): 152–58.

_____. 1997. "Structure Mapping in Analogy and Similarity." *American Psychologist* 52/1 (January): 45–56.

Johnson, M. 1987. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason.* Chicago: University of Chicago Press.

Lakoff, G. 1990. "The Invariance Hypothesis: Is Abstract Reason Based on Image-schemas?" *Cognitive Linguistics* 1/1: 39–74.

Lakoff, G., and M. Johnson. 1980. *Metaphors We Live By.* Chicago: University of Chicago Press.

Proust, M. 1981. *Remembrance of Things Past, Vol. 1, Swann's Way.* Trans. C. K. Scott Moncrieff and T. Kilmartin. New York: Vintage Books.

Tovey, D. F. 1938. "The Main Stream of Music." *Proceedings of the British Academy* 24: 113–41.

Turner, M. 1990. "Aspects of the Invariance Hypothesis.' *Cognitive Linguistics* 1/2: 247–55.

_____. 1996. *The Literary Mind.* New York: Oxford University Press.

Zbikowski, L. M. 1998. "Metaphor and Music Theory: Reflections from Cognitive Science." *Music Theory Online* 4/1 (January).

_____. 2002. *Conceptualizing Music: Cognitive Structure, Theory, and Analysis*. New York: Oxford University Press.