

Chapter 15

Music, language, and multimodal metaphor

Lawrence M. Zbikowski

Abstract

This chapter considers the topic of multimodal metaphor from the perspective of cross-domain mappings between the musical and the linguistic domains. Beginning with an example of what musicians call “text painting” (in which music is used to “paint” an image related to the text of a vocal work), I explore the different ways music and language structure thought. Examples of musical passages from Palestrina, Biber, Bach, Schubert, and Jerome Kern are used to demonstrate how music contributes to meaning construction and thus may serve as a source domain for a multimodal metaphor. I conclude with a brief discussion of how conceptual blending theory can be used for the analysis of text-music relations, and the multimodal metaphors that may result.

Keywords: music, text-painting, conceptual blending, popular song, musical grammar, Bach, Schubert

1. Introduction

Giving voice to an idea that took a number of forms in his later work, Ludwig Wittgenstein, near the end of his *Philosophical Investigations*, wrote, “Understanding a sentence is much more akin to understanding a theme in music than one may think” (Wittgenstein 2001: frag. 527). The close relationship between language and music suggested by Wittgenstein’s observation is borne out by similarities between the two: both are unique to the human species, both unfold over time, both have syntactic properties, and both make use of sound. There are also, of course, notable differences: musical meaning is on the whole much less precise than linguistic meaning; music often involves simultaneous events, where language does not; and there is more of a sense of play in ordinary music than there is in ordinary

language. On the one hand, these differences suggest that language and music belong to two different conceptual domains. On the other hand, the similarities between the two suggest that language and music may recruit some of the same cognitive resources, and that structure from one domain may be readily mapped to the other to create meaning. Understanding a sentence is like understanding a musical theme because both language and music offer possibilities for constructing meaning, possibilities that can be exploited through multimodal metaphors.

As an example, consider the passage from the Credo of Giovanni Pierluigi da Palestrina's *Pope Marcellus Mass* (printed 1567) given in figure 1.¹ The text Palestrina sets here is "Qui propter nos homines et propter nostram salutem descendit de caelis" ("Who for us men, and for our salvation, came down from heaven"). With the first statement of the word "descendit," each voice begins a scalar descent. Christ's descent from heaven is thus represented with a cascading fall through musical space, a series of overlapping movements "down" the musical scale. This representation exploits the common construal of musical pitches as situated in vertical space, a construal that follows from the characterization of pitches as "high" or "low" with respect to one another.

The image shows a musical score for six voices: Cantus, Altus, Tenor I, Tenor II, Bassus I, and Bassus II. Each voice part has a staff with a treble or bass clef and a key signature of one sharp (F#). The lyrics are: "tem. de - scen - dit de cae - lis." The Cantus part starts on a high note and descends stepwise. The other voices enter at different points, creating a cascading effect of descending lines. The Tenor I part has a fermata over the word "tem".

Figure 1. Giovanni Pierluigi da Palestrina, Credo of the Pope Marcellus Mass, mm. 53-58.

Although this characterization seems quite natural, it is actually rather arbitrary. For instance, when you play "higher" notes on the piano, you move your hand to the right; when you play "higher" notes on a cello, you move your hand down toward the ground.² Describing musical pitches in terms of "high" and "low" is in fact a product of mapping structure from one domain of knowledge (relationships among points in vertical space) onto another (relationships among musical pitches). From the perspective provided by the contemporary theory of metaphor (Lakoff 1993), this mapping relies on the conceptual metaphor PITCH RELATIONSHIPS ARE RELATIONSHIPS IN PHYSICAL SPACE. Language, of course, facilitates this mapping. In that language belongs to one mode (that of either spoken or written signs) and music to another, the metaphorical description of pitches as "high" or "low" is multimodal in at least a minimal sense. If, however, we take a closer look at Palestrina's compositional choices, a somewhat more textured view of multimodal metaphors that involve music emerges. Were Palestrina interested only in portraying a move from "high" to "low," he could have used a single falling interval rather than his stepwise descending scale. Using the resources offered by six voices, he also reinforced and inflected his sonic image of descent through a series of multiple, and subtly different, descending gestures, all of which culminate in the important cadential arrival of measure 58. What Palestrina gives us, then, is the *sound* of descent, realized as an orderly, stately process. Contrast his setting of "descendit" with that of Heinrich Biber at the corresponding point in the Credo from Biber's *Missa Christi resurgentis*, written a little over a hundred years after Palestrina's mass. As shown in figure 2, the path sketched by Biber is much more involved, comprising fifteen notes and proceeding through a series of twisting turns before reaching its goal. Palestrina, for his part, takes only eight notes, and never changes direction.³ The sonic image of descent offered by Biber is consequently quite different than that offered by Palestrina. Biber's descent

The image shows two musical staves. The top staff is for Heinrich Biber, showing a complex melodic line with many notes and some chromaticism. The bottom staff is for Giovanni Pierluigi da Palestrina, showing a much simpler, stepwise descending line. Both staves have the lyrics "de - scen - dit de cae - lis" written below them.

Figure 2. Comparison of melodic passages from Heinrich Biber, Credo of the *Missa Christi resurgentis*, mm. 51-57 and Giovanni Pierluigi da Palestrina, Credo of the *Pope Marcellus Mass*, mm. 53-55.

is a leisurely, almost caressing, affair in which the journey is at least as important as the goal. Indeed, when Biber approaches the cadence that articulates the end of this section of the Credo he leaves behind the melodic material of figure 2 and adopts a much more proclamatory style, effectively bracketing off the sensation of descent from its goal.

Palestrina's and Biber's settings of "descendit" demonstrate, in a concise way, the different resources offered by language and music for the construction of multimodal metaphors. Language gives access to a rich network of conceptual frameworks. "High" and "low," for instance, are used not only for orientation in physical space, but also as evaluative terms: Christ's descent is from heaven, where everything is good, to earth, where it is not, a perspective that relies on the conceptual metaphor STATE OF BEING IS ORIENTATION IN VERTICAL SPACE.⁴ As another example of the conceptual frameworks to which language offers access, the continuum of vertical space can be divided up into a series of points with distances between them. When these points are mapped onto musical pitches (through the conceptual metaphor PITCH RELATIONSHIPS ARE RELATIONSHIPS IN PHYSICAL SPACE) it becomes possible to measure the "distance" between any two notes. I would emphasize that neither of these conceptual frameworks has immediate relevance for Palestrina's or Biber's setting of "descendit," yet both are accessible through the conceptual networks to which the "high" and "low" of musical space is connected. Music, for its part, infuses the process of meaning construction with a crucial dynamic aspect. It is not simply the concept of descent through vertical space that we draw from the musical domains set up by Palestrina and Biber but particular *kinds* of descent, each with its own texture and shape. Where language provides us with only the bare prompt of the word "descendit," a notion we fill out with our own experiences with the process of descent, music actively sketches different sorts of descents through carefully arranged sonic materials.

In this chapter I want to explore in greater detail multimodal metaphors that involve language and music from both theoretical and practical perspectives. The theoretical perspective, which is developed in the first main section that follows, concerns the different functions of language and music in human culture. I should note that in this section, and indeed throughout this chapter, I shall be concerned with the ways musical materials actually give rise to meaning rather than the meanings with which they may become associated.⁵ The practical perspective, which is developed in the second main section, begins with a treatment of my opening examples in terms of the theoretical perspective for which I have argued and expands the discussion to consider relationships between text and music in greater detail. In a con-

cluding section I shall return to Wittgenstein's observations about understanding in language and music, and consider what multimodal metaphors can tell us about how we understand language and music.

2. Function and structure in language and music

As was noted above, there are numerous similarities between language and music, and there are any number of cultural practices that blur the boundaries between the two modes of communication (Boiles 1967; Fornäs 2003). It is still the case, however, that we do not live our lives inside the equivalent of a grand opera or Broadway musical, with every utterance sung and every action accompanied by an orchestra. So why is it that human cultures have developed both language and music? Although there is any number of ways to approach this question, I take the position that language and music have different functions within human culture. In doing so I am influenced by the work of the developmental psychologist Michael Tomasello, who situates the emergence of language in our species within the broader development of human culture. In Tomasello's view, the primary function of language is to direct the attention of another person to objects or concepts within a shared referential frame (Tomasello 1999: chap. 5). I would argue that music is similarly part of a cultural framework unique to our species, but one whose primary function is to represent through patterned sound various dynamic processes that are common in human experience. Chief among these dynamic processes are those associated with the emotions (which, following recent work by Antonio Damasio, can be construed as sequences of physiological and psychological events that subtend feelings [Damasio 1999, 2003]) and the movements of bodies – including our own – through space.

The difference in function between these two modes of communication is matched by a difference in the forms through which the functions are realized. In the case of language, this is accomplished through symbolic units that correlate with functions such as those represented by nouns and verbs, as well as with the many other parts of speech recognized by grammarians (Croft 2001: chap. 2). Through the use of these symbolic units we can direct the attention of another person to objects or concepts within a shared referential frame.

In the case of music, the basic formal unit is what I call a sonic analog, which represents through patterned sound the central features of some dynamic process.⁶ "Descent," for instance, is one such dynamic process: Palestrina, in the example shown above, provides one sonic analog for this

process; Biber, in the Credo from his mass, provides another. I would hasten to add that neither musical passage *stands for* descent in a direct or unequivocal fashion. Instead, the phenomenal properties of these passages are such that they can serve as analogs for the process of descent.⁷

Given this perspective, it should be clear that mappings between the domains of language and music will involve structures that are fundamentally different in kind. Language tends to focus on objects (whether real or imagined) and relationships between objects. Language can direct our attention to a process (that is, the noun “descent” picks out a dynamic process that involves a traversal of space), but it is less common for language to *embody* such a process. When it does – when, for instance, we imitate the sound of a horse’s step with the words “clip-clop, clip-clop” – language starts to become more like music. Music, for its part, does not tend to be involved with the rich symbolic systems typical of language.⁸ In those cases where music does exploit this sort of symbolic system – as, for instance, through musical topics of the type employed by Mozart and Haydn (Ratner 1980; Allanbrook 1983; Agawu 1991, 1999) – its dynamic aspect tends to recede in importance. In sum, then, mappings from language to music will tend to focus on static aspects of the musical domain; mappings from music to language will draw out the dynamic aspects of the domain of language.

Although any attempt to determine a crisp boundary for what counts as language or what counts as music may be an endeavor destined to generate more heat than light, it seems clear that, at least in their characteristic usage in the contemporary world, language and music have different functions. While the range of language functions is broad, primary among these is the use of symbolic tokens to direct the attention of another person to objects and relations within a shared referential frame. Music, by contrast, provides sonic analogs for a wide range of dynamic processes that are marked in human experience, especially those associated with the regulation of emotions. Multimodal metaphors that involve language and music draw on both of these resources, as the analyses in the next section shall make clear.

3. Text Painting

Because so much of our communication is done through language, the contribution of non-linguistic forms of communication to the construction of meaning often goes unnoticed. However, as work on the gesture that accompanies speech has shown (Goldin-Meadow 2003a; Goldin-Meadow 2003b; Kendon 2004; McNeill 1992, 2005; Müller and Cienki this volume; Mittel-

berg and Waugh this volume), non-linguistic forms of communication are an important way humans shape their thought. Evidence for a similar role for prosody has been provided by a recent set of experiments that demonstrated the use of pitch inflection as an analog for motion through space, and rate of delivery as an analog for the speed of an object (Shintel, Nusbaum, and Okrent 2006). With regard to the latter, when subjects were asked to characterize the motion of a dot that moved horizontally across a screen at one of two speeds by saying either “It is going left” or “It is going right,” they spoke more rapidly if the dot was moving quickly than if it was moving slowly. Moreover, listeners were able to make accurate judgments about the speed of the dot when they heard recordings of these statements.

In the case of multimodal metaphors that involve language and music, a consequence of the emphasis on language has been that mappings from language to music have received the most attention (Zbikowski 2008). Mappings from music to language – which function, as do gesture and prosody, to shape thought – have been, by comparison, rather neglected (although see Lidov 2005 for a recent corrective). The compositional technique of text painting provides an opportunity to explore both ways of mapping in more detail, and to develop a better understanding of multimodal metaphors that involve music.

The basic idea of text painting is simple enough. When a particularly strong or compelling image occurs in the text for a musical work, the composer writes the accompanying music to suggest, or “paint,” the image. Thus if the text mentions a galloping horse, the music coincident with the text might imitate the sound and action of a horse proceeding at full speed. The passages from Palestrina’s and Biber’s masses that I discussed in my introductory comments are typical examples of text painting, but with one illuminating characteristic: there is in fact nothing in Palestrina’s or Biber’s music that imitates the sound of an actual descent. Indeed, the sounds associated with descents have little in common with a descending scale, and in some cases – for instance, walking down a hill – descent may be virtually soundless. Text painting is, in consequence, not so much about imitating some naturally-occurring sound as it is about providing sonic analogs for various dynamic processes. In Biber’s and Palestrina’s masses, the music provides a sonic analog for the dynamic process associated with descent (correlating descending pitch with a decrease in potential energy), and our understanding of this analog is structured by the accompanying text.

In terms of the contemporary theory of metaphor, text painting occurs when concepts prompted by the text for a musical work (which serves as the source domain) are mapped on to a series of musical events (which serves as

the target domain). This mapping structures our understanding of the musical domain: we hear the sounds as “descending.” But as I suggested in my introductory comments, the sonic analog provided by the music also shapes our understanding of the text, for the music gives the delivery of the words a specific contour and duration.⁹ Again, given our tendency to give priority to linguistic domains, the notion that the seemingly indefinite and nonconceptual domain of music could be used to structure thought may seem at best little more than a passing curiosity, and at worst downright nonsense. But consider three situations: “descendit” spoken; “descendit” sung by Palustrina’s singers; and “descendit” sung by Biber’s singers. If there are any differences between these three utterances, they come from the structure music can impose on language.

Two further examples of text painting can help to elaborate this point. The first of these comes from the fourth movement of Johann Sebastian Bach’s Advent cantata “Nun komm der Heiden Heiland” (BWV 61). The text and music for each of the three preceding movements of the cantata have all focused on the coming of Christ (as befits the Advent theme). In the fourth movement Christ is suddenly before us, speaking words from the third chapter of Revelation: “Behold, I stand at the door, and knock. If any man hear my voice, and open the door, I will come in to him, and will sup with him, and he with me.” Bach sets this text as an accompanied recitative for baritone, with the strings playing pizzicato throughout; a shorthand version of the string parts and solo bass melody for the first four measures is given in figure 3.

The image shows a musical score for measures 1-4 of the fourth movement of J. S. Bach's cantata "Nun komm der Heiden Heiland" (BWV 61). The score is written for a baritone voice and a piano accompaniment. The bass line (bottom staff) features a melodic line with a descending contour, while the piano accompaniment (top and middle staves) consists of a rhythmic pattern of sixteenth notes. The lyrics are: "Sie-he, sic-he. ich ste-he vor der Tur und klo - pfe an, und klo-pfe an." The tempo marking is "♩=54".

Figure 3. Measures 1-4 of the fourth movement (Recitativo) from J. S. Bach’s cantata “Nun komm der Heiden Heiland” (BWV 61).

Bach’s text painting centers on the words “und klopfe an” – that is, “and knock.” Bach uses three compositional techniques to paint this activity. First, he summons the repetitions we associate with the act of knocking by repeating the words, and by using three notes to set the first syllable of the initial “klopfe” (a device called a melisma) such that the articulation of the

syllable is repeated. Second, he uses staccato marks on the three notes of the melisma, which place silences between these notes; these silences are similar to those that fall between knocks on a door. Third, he sets the words with a broken chord (or arpeggio). This places a kind of distance between each successive note but also allows us to hear all as belonging to a single connected gesture.

Again, as an out-and-out imitation of the act of knocking Bach’s setting of “und klopfe an” leaves something to be desired. Knocks are usually unpitched, for instance, but Bach gives us different pitches for each blow. Knocking is not usually accompanied, but here we have pizzicato strings pulsing in the background.¹⁰ The reason we hear these musical events as knocking (to the extent that we do) is that our understanding of them is structured by mapping concepts from language onto music. Again, once we hear these musical events as the sound of knocking our understanding of the text starts to change, for the music creates a rather specific analog for the dynamic process of knocking on a door. This sonic analog relies on the rhythmic and pitch resources of music: the delivery of “und klopfe an” is at first halting and then (when it is repeated) hurried; the rhythmic structure of the passage is coordinated with a notable expansion of pitch space (from the span of B3 to D#3 in measures 1-2 to the span of C4 to B2 in the first part of measure 3) which is then compressed by the final F#3-B3-F#3-G3. These resources in turn shape our understanding of the text. We not only know that someone is knocking at the door, but *how* they are knocking: first tentatively, and then with more urgency.

My final example of text painting is from Franz Schubert’s song “Gretchen am Spinnrade,” which takes as its text a scene from Goethe’s *Faust*. In the scene we overhear Gretchen as, alone in her room, she describes her love for – or perhaps enchantment by – Faust. The song begins, as shown in figure 4, with the briefest of introductions by the piano, a swirling sixteenth-note figure that circles around a D minor chord, a few sparse and repetitive bass notes sounding beneath. This pattern, or some version of it, continues throughout the entire 120 measures of the song, with but one interruption (to which we shall return in a moment). At first glance, the text (the first lines of which translate as “My peace is gone, / My heart is heavy, / I will find it never / and never more”) may seem to have little to do with this monotonous accompaniment. The link is provided by the title: this is Gretchen at the spinning wheel. Schubert’s accompaniment is, of course, meant to evoke the sound of the wheel in action, with the swirling sixteenth notes summoning the wheel itself and the repetitive, off-beat accents in the middle voice representing the clack of the bobbin, but for modern listeners

this sonic image will not typically resolve itself until we have mapped conceptual structure from the domain of the text onto the domain of the music. The text thus serves a function similar to what Roland Barthes has called “anchoring,” rendering transparent an aural image that might otherwise remain opaque.

The image shows a musical score for three systems of measures 1-13 from Franz Schubert's "Gretchen am Spinnrade." Each system consists of a vocal line (soprano) and a piano accompaniment. The first system (measures 1-4) is titled "Nicht zu geschwind (M.M. ♩ = 72)" and contains the lyrics "Mei-ne Ruh- ist hin, mein Herz ist". The second system (measures 5-8) contains the lyrics "schwer, ich fin-de, ich fin-de sie nim-mer und". The third system (measures 9-13) contains the lyrics "nim-mer-mehr. Wo ich". The piano accompaniment features a constant sixteenth-note pattern in the right hand, with a dynamic marking of *pp* (pianissimo) at the beginning and *cresc.* (crescendo) in the middle. The vocal line is in a simple, lyrical style.

Figure 4. Measures 1-13 from Franz Schubert's "Gretchen am Spinnrade," op. 2 (D. 118).

Again, Schubert's accompaniment does not, in any direct way, imitate the sound of a spinning wheel (which is unpitched) although it does provide a surprisingly accurate analog for the act of spinning. Typical treadle speeds start at about 60 treadles per minute, with the main wheel turning around once with each push of the treadle.¹¹ Each complete pattern of sixteenth notes in Schubert's accompaniment (with two complete patterns per measure) takes a bit less than a second to complete at standard performance tempos, meaning that Schubert's "wheel" spins at approximately the same speed

as does Gretchen's. And, as does Gretchen's wheel, Schubert's continues uninterrupted until, midway through the song, Gretchen seems to forget herself and the accompaniment momentarily breaks off. The resumption of the musical pattern – and, presumably, Gretchen's spinning – is at first halting, getting under way only with the return to the music shown in figure 4.

In my other examples of text painting, the mappings between language and music were more or less focused: a particular musical passage provided a sonic analog for descent, or for the act of knocking. In "Gretchen am Spinnrade" the mappings are rather less focused: the correlation is not between a word or cluster of words and a particular musical passage, but between the situation described in the title and the accompaniment pattern that permeates the song. As a result, the music informs our understanding of not just one word from the text but the text as a whole. Through the music we can hear Gretchen's obsession with Faust (in the relentless patterns of the accompaniment) and sense its fevered intensity: over the course of the song Schubert makes sparing use of the sort of normative harmonic syntax used to suggest progress, and relies instead on quicksilver gestures toward various harmonic centers indicative of a mind that cannot settle down to anything.

The curious situation evident in this last case of text painting, where a sonic analog informs our understanding of an entire scene rather than just a single word, points to an interesting feature of multimodal metaphors that involve language and music, and that can be illuminated by a brief consideration of directionality in metaphor. As has often been noted, metaphors are directional: the statement "The hippopotamus is a ballerina" is rather different from the statement "The ballerina is a hippopotamus." In the first case, attributes associated with a ballerina are mapped on to the hippopotamus; in the second case, attributes associated with a hippopotamus are mapped on to the ballerina. One relatively straightforward explanation for the obvious differences between these two mappings views metaphor as a special case of analogy, in which the correlated domains are conceptually distant from one another (Holyoak and Thagard 1995: 213). From this perspective, the directionality of metaphor can be viewed as a consequence of mapping conceptual structure from a source domain to a target domain when there is relatively little conceptual overlap between the two; when the mapping is reversed, the large amount of new information introduced creates a second metaphor markedly different from the first. In the case of linguistic metaphors, the more abstract structure of the correlated domains is nonetheless retained: objects are mapped to objects ("ballerina" to "hippopotamus") and relations to relations ("dances" to "lumbers"). Given the account of the different functions of language and music that I have offered, however, even this aspect of

the lyrics, the mental space set up by the music, and then the blended space that results from their combination.

Fields's lyrics open with a characteristic space builder: "Someday."¹³ Here, the space builder establishes a mental space focused on a future state of affairs rather than on the present. The followings lines fill out the picture: the speaker, beset with rather dire circumstances ("When I'm awfully low / And the world is cold") will be comforted by the remembrance of the object of his affections and, more specifically, by the way she looks on this particular night ("I will feel a glow just thinking of you / And the way you look tonight"). As is hinted at by the transformative effect of the appearance of the beloved (an effect confirmed by the second verse's "There is nothing for me but to love you"), what is involved here is not simply a kind of passive looking, with one person gazing on another, but an intimacy of association that has both power and depth. The mental space established by the first verse thus develops into a scene in which what is of moment is not some future opportunity to look back to the present as a golden past but the centrality of "the way you look tonight" to a highly charged romantic relationship. The "Someday" space builder is thus somewhat misleading: what is important is not the future but the present.

The melody for the song also begins with a space builder, but in this case the falling fifth A5-D4. Although these pitches could be understood in a variety of ways, the simplest interpretation (and one supported by the opening D major chord) is as the fifth and first notes in a D scale. In the music that follows the registral space between D4 and A5 is filled in by a sequence of arch-like figures that flesh out the musical topography with notes from the key of D major (rather than, for instance, D minor). These figures ultimately move past A5 (in measure 7) and arrive on the high note of the melody (D5) of measure 9. This arrival coincides with a return to the rhythm of the opening gesture, but with the A5-D4 falling fifth replaced by a D5-D4 descending octave. It is worth noting that while the D4 of measure 2 and the D4 of measure 10 are the same pitch, their context is quite different: the registral space above the D4 of measure 10 has been expanded through the sequential figures of measures 3-8, and the temporal space between the two whole notes has been filled in by the moving quarter-notes in these same measures. Measure 10 is followed by a final passage that owes something to the compositional strategies that filled out the musical space in measures 3-8, but which now lead directly to the D4 that concludes the melody of the A section in measure 13. From a musical perspective, then, the song opens with a space builder roughly equivalent to the "Someday" of the words, after which

the space is filled out – with notes instead of words – to solidify and stabilize the conceptual realm prompted by the space builder.

The melodic process that leads to the arrival on D4 in measure 13 is reinforced by rhythmic and harmonic processes that are to some extent independent of the melody.¹⁴ Over the course of measures 1-8 the number of shorter-duration notes increases until it reaches its maximum density in measures 7-8. This is followed (as has already been noted) by the whole notes of measures 9-10. The contrast between shorter- and longer-duration notes is then revisited in a more orderly fashion, with the quarter notes of measure 11, half notes of measure 12, and whole note of measure 13 (tied over into the quarter note of measure 14). While the harmonies used by Kern – a four-measure pattern repeated (with slight variation) four times – are quite typical in American popular music (what musicians would call a I-VI-II-V pattern), two details contribute to the overall dynamic shape of the A section. First, although the D major harmony of measure 5 represents a return to the opening harmony, the music is kept moving by the melodic sequence that accompanies this harmony. Second, although D major returns once more in measure 9 (and supports long notes that recall the long notes of the opening measures), Kern destabilizes the chord by turning it into a dominant seventh. The overall effect of all of these processes – melodic, rhythmic, and harmonic – is to make measures 11-13 a goal for the A section, a goal whose culmination is the D4 of measure 13, the note which correlates with the arrival on the last syllable of "The way you look *tonight*."

Essential features of the mental spaces set up by the lyrics and the music for the first verse of "The Way You Look Tonight" are given in the conceptual integration network diagrammed in figure 6, with each mental space represented by a circle. Where the mental space set up by the text is concerned primarily with objects and relations (namely, the appearance of the beloved and its importance for "romance") the mental space set up by the music is concerned with a set of coordinated processes that lead to the final phrase in the section. When these two spaces are correlated with one another (as they are in the song) aspects of their structure are projected into the blended space to yield a *dynamic* representation of the development of an intimacy of association. Guiding the integration of these concepts is the generic space (represented in the top circle of the diagram), which defines the core cross-space mapping and basic topography for the network. The generic space, which reflects the insights captured in the invariance principle (Lakoff 1990; Turner 1990), is centered on the idea that focused attention is a form of intimacy. This sort of attention is behind the fixed gaze of the lover, and it

is also behind our discomfort when confronted with the unbending stare of a stranger.

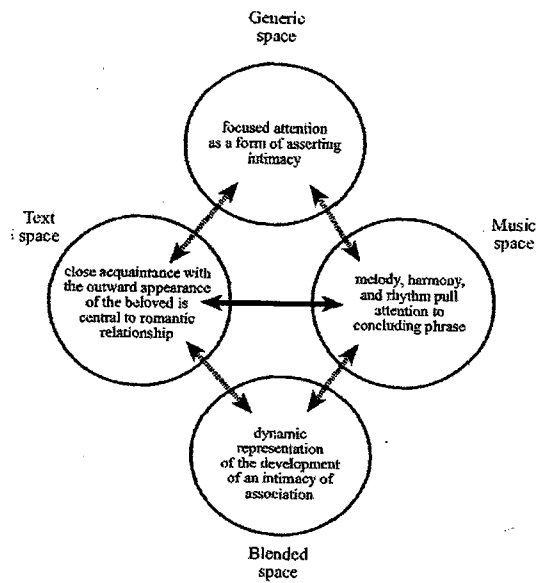


Figure 6. Conceptual integration network for the first verse of "The Way You Look Tonight."

The development of this sort of intimacy was at the heart of the scene from the 1935 movie *Swing Time* (starring Fred Astaire and Ginger Rogers) where "The Way You Look Tonight" first appeared. In this scene Astaire is trying to win Rogers over: he has managed to get in to her hotel suite, and even though she has repeatedly spurned him, gives it his best effort by singing this song ("accompanying" himself at the piano). Rogers, for her part, has locked herself in her room and has started shampooing her hair. As the song unfolds, however, she emerges from her room in a bathrobe with hair lathered, first smiling on Astaire and then soundlessly walking over to stand behind him at the piano, where she rests her hand on his shoulder. This gesture coincides with Astaire's arrival at the music of measure 61 (the penultimate statement of "The way you look tonight"), and in response to it he sings the words one more time, turning to look into Rogers' eyes. At the conclusion of the song his gaze becomes quizzical as he notices the lather on her head and she, observing this change, turns to a mirror and discovers her appearance. This comedic moment is, of course, a play on "the way you look

tonight," but Astaire and Rogers have just shared an encounter typical of a much more intimate relationship than they have enjoyed thus far. This intimacy is one that is worked out not only in the first verse of Kern and Field's song, but over its entire course. The second verse moves from generalities of appearance to specifics ("With your smile so warm / And your cheek so soft"), the bridge (with its music momentarily suspending the process enacted by the first two A sections) adds the detail of "that laugh that wrinkles your nose," and the final verse finds the singer speaking directly to the beloved: "Never, never change."

From the perspective of conceptual blending theory, then, in some cases words and music will prompt the construction of two independent but correlated mental spaces. Both of these spaces contribute structure to a third mental space, in which concepts drawn from each of these two input spaces are blended. This new space typically serves as a site for the imagination. For the conceptual blend created by the words and music for "The Way You Look Tonight," we might well imagine that the intimacy established between the lovers is one that would lead them to dance together, or to exchange loving words, or perhaps just to stare into each other's eyes. None of these possibilities concludes the scene from *Swing Time* – Rogers, aghast at her appearance, rushes back to her boudoir – but subsequent scenes do make clear that Rogers' and Astaire's characters are now a pair.

The mappings associated with a conceptual blend of this sort are different from those associated with a metaphor in two important ways. First, blending typically involves highly fluid and thoroughly pragmatic mental spaces rather than established domains. Metaphorical mappings often yield systems of metaphor; while blends may exploit such systems, they may also destabilize them by extending the system in novel ways, and thus push against the boundaries of the domain. Second, rather than one domain (the source) providing structure for the other (the target) – a mapping that gives rise to the directionality of metaphor noted above – in a blend correlated spaces each contribute to structure that is mapped onto the blend. But whether metaphor or conceptual blending is involved, I hope I have demonstrated the resources for meaning construction provided by these two different modalities. Most readers will find mapping from language to music to be simple enough, not the least because language is the primary means most humans use to structure their understanding of the world. Mapping from music to language might seem a stranger alternative, but I believe it is a real possibility, especially when language is for some reason ambiguous (as it is in the opening of Schubert's "Gretchen am Spinnrade"). Additional evidence that music can serve as a proper conceptual domain, and thus have at least the *potential* to

structure our understanding of language, comes from instances of conceptual blending where music provides one of the input spaces for a conceptual blend. (I discuss further instances in Zbikowski 2002: chaps. 2 and 6.)

5. Conclusion

Although Wittgenstein believed that how we understand music provided important insights into how we understand anything in general – and language in particular – he himself struggled with what it meant to understand music. In a fragment that dates from 1948 he wrote, “If I now ask ‘So what do I actually experience when I hear this theme and understand what I hear?’ – nothing occurs to me by way of reply except trivialities. Images, sensations of movement, recollections and such like” (Wittgenstein 1980: 69e-70e). Multimodal metaphors that involve language and music have the potential to provide crucial insights into this question, for such metaphors make clear the different functions of language and music in human culture, and the different ways they construct meaning. Language gives us the means to represent symbolically objects and relations, and through these representations we can direct the attention of another person to things within a shared referential frame. Music, by contrast, provides us with sonic analogs for dynamic processes, processes that include movement through space (such as descent), physical gestures (like knocking), and emotional states (such as obsession or the development of intimacy). A place to begin understanding how we understand music is with such sonic analogs.

Although the established models of cross-domain mapping can tell us much about the different contributions language and music make to multimodal metaphors, the methodology of conceptual blending, and the somewhat more fluid construct of a mental space basic to this methodology, offers one way to capture the unique contribution of each mode of communication to the process of meaning construction. Conceptual blending raises at least as many questions as it answers, but it has the potential to place language and music on an equal footing so that we may better understand the contribution of each to multimodal metaphor.

As Wittgenstein suggested, understanding a sentence is much more akin to understanding a musical theme than we might first think. This is not to say that language and music accomplish their ends by the same means – I have in fact argued just the contrary – but that they draw on some of the same cognitive resources for constructing meaning. A key to understanding how this is possible is offered by multimodal metaphors that involve both

language and music, for such metaphors give us compelling insights into the rich and varied world of meaning within which humans dwell.

Notes

1. I discuss this passage at greater length in Zbikowski (2002: chap. 2).
2. Further discussion of ways to characterize relationships between musical pitches can be found in Zbikowski (2002: chap. 2) and Ashley (2004).
3. For the sake of concision I have given only one of Biber’s melodic strands. At this point in the Credo there are nine vocal parts, and together with the instrumental ensemble the melodic contour shown in example 2 is replicated no fewer than four times (in most cases with pairs of voices). The temporal span of Biber’s setting of “descendit de caelis” is also significantly longer than that of Palestrina: Biber’s setting runs to forty-five seconds, where Palestrina’s is only about twenty seconds long.
4. This conceptual metaphor is a variant of STATES ARE LOCATIONS discussed by Lakoff and Turner (1989).
5. With regard to the associational meanings of music, it is worth noting that one of the most well-known depictions of the pounding of horse’s hooves in American popular culture – the use of a portion of Gioacchino Rossini’s overture to *Guillaume Tell* for the theme music for the radio and television series *The Lone Ranger* – does not have any associations with horses in its original context, but with general ideas about victory.
6. As I use the term, “representation” does not have to entail a full semiotic system. This perspective is similar to that adopted by Naomi Cumming (2000: chap. 3). A dynamic process may be provisionally defined as a coherent sequence of phenomena that is distributed over time and typified by parametric modulation or change.
7. Sonic analogs are akin to Charles Peirce’s notion of an icon, and in particular to the form of an icon he called a diagram, “which represent the relations ... of the parts of one thing by analogous relations in their own parts” (Peirce 1960, 1:277).
8. I draw my perspective on the unique symbolic aspect of language from Deacon (1997, 2003).
9. The rough equivalent in prosody would combine the two parameters studied by Shintel and her colleagues, yielding something like “It is going doooooowwwwnn” spoken with a descending inflection.
10. A few writers have gone so far as to interpret the steady plucking of the orchestra in this movement as a further embodiment of knocking, but this seems something of a stretch. Not only are the attack points too widely spaced to sound much like knocking but the effect is far too persistent, more like Edgar Allen Poe’s telltale heart than a summons from the Savior.
11. This information was gleaned from <http://www.hjsstudio.com/espinner.html> (accessed 26 September 2006).

12. For an overview of the theory see Fauconnier and Turner (1998, 2002). For a review of the latter, see Forceville (2004).
13. Fauconnier describes a space builder as follows: "Linguistic expressions will typically establish new spaces, elements within them, and relations holding between the elements. I shall call *space-builders* expressions that may establish a new space or refer back to one already introduced in the discourse" (1994: 17).
14. Music scholars often analyze music in terms of three primary parameters: melody, harmony, and rhythm. In the same way that the very notion of "music" varies broadly across cultural practices, the manifestation of these parameters is not always obvious or unequivocal. In "The Way You Look Tonight," however, both the harmony (indicated by the chord symbols above the staff) and the rhythmic frame (indicated by the notated durations of pitches and by barlines) are relatively clear-cut.

References

- Agawu, Victor Kofi
 1999 The challenge of semiotics. In *Rethinking Music*, Nicholas Cook and Mark Everist (eds.), 138–60. Oxford: Oxford University Press.
 1991 *Playing with Signs: A Semiotic Interpretation of Classical Music*. Princeton: Princeton University Press.
- Allanbrook, Wye Jamison
 1983 *Rhythmic Gesture in Mozart: Le Nozze di Figaro and Don Giovanni*. Chicago: University of Chicago Press.
- Ashley, Richard
 2004 Musical pitch space across modalities: Spatial and other mappings through language and culture. In *Proceedings of the 8th International Conference on Music Perception and Cognition*, Scott D. Lipscomb, Richard Ashley, Robert O. Gjerdingen, and Peter Webster (eds.), 64–67. Adelaide, Australia: Causal Productions.
- Boiles, Charles L.
 1967 Tepehua thought-song: A case of semantic signaling. *Ethnomusicology* 11 (3): 267–92.
- Cook, Nicholas
 2001 Theorizing musical meaning. *Music Theory Spectrum* 23: 170–95.
- Croft, William
 2001 *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.
- Cumming, Naomi
 2000 *The Sonic Self: Musical Subjectivity and Signification*. Bloomington: Indiana University Press.
- Damasio, Antonio R.
 1999 *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace.
 2003 *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*. Orlando, FL: Harcourt.
- Deacon, Terrence W.
 1997 *The Symbolic Species: The Co-Evolution of Language and the Brain*. New York: W.W. Norton.
 2003 Universal grammar and semiotic constraints. In *Language Evolution*, Morten H. Christiansen, and Simon Kirby (eds.), 111–139. Oxford: Oxford University Press.
- Fauconnier, Gilles
 1994 *Mental Spaces: Aspects of Meaning Construction in Natural Language*. 2nd ed. With a foreword by George Lakoff and Eve Sweetser. Cambridge: Cambridge University Press.
- Fauconnier, Gilles, and Mark Turner
 1998 Conceptual integration networks. *Cognitive Science* 22: 133–87.
 2002 *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.
- Forceville, Charles
 2004 Review of Gilles Fauconnier and Mark Turner, *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. *Metaphor and Symbol* 19: 83–89.
- Fornäs, Johan
 2003 The words of music. *Popular Music and Society* 26: 37–51.
- Goldin-Meadow, Susan
 2003a *Hearing Gesture: How Our Hands Help Us to Think*. Cambridge, MA: Harvard University Press.
 2003b *The Resilience of Language: What Gesture Creation in Deaf Children Can Tell Us About How All Children Learn Language*. New York: Psychology Press.
- Holyoak, Keith J., and Paul Thagard
 1995 *Mental Leaps: Analogy in Creative Thought*. Cambridge, MA: MIT Press.
- Keller, Hans
 1994 The musical analysis of music. In *Essays on Music*, Christopher Wintle (ed.), 126–28. Cambridge: Cambridge University Press.
- Kendon, Adam
 2004 *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Lakoff, George
 1990 The invariance hypothesis: Is abstract reason based on image-schemas? *Cognitive Linguistics* 1: 39–74.

- 1993 The contemporary theory of metaphor. In *Metaphor and Thought*, 2nd ed. Andrew Ortony (ed.), 202–51. Cambridge: Cambridge University Press.
- Lakoff, George, and Mark Turner
1989 *More Than Cool Reason: A Field Guide to Poetic Metaphor*. Chicago: University of Chicago Press.
- Lidov, David
2005 *Is Language a Music? Writings on Musical Form and Signification*. Bloomington: University of Indiana Press.
- McNeill, David
1992 *Hand and Mind: What Gestures Reveal About Thought*. Chicago: University of Chicago Press.
2005 *Gesture and Thought*. Chicago: University of Chicago Press.
- Mittelberg, Irene, and Linda R. Waugh
this vol. Multimodal figures of thought: A cognitive-semiotic approach to metaphor and metonymy in co-speech gesture.
- Müller, Cornelia, and Alan Cienki
this vol. Words, gestures and beyond: Forms of multimodal metaphor in the use of spoken language.
- Peirce, Charles Sanders
1960 *Collected Papers of Charles Sanders Peirce*, Charles Hartshorne and Paul Weiss (eds.). Cambridge, MA: The Belknap Press of Harvard University Press.
- Ratner, Leonard G.
1980 *Classic Music: Expression, Form, and Style*. New York: Schirmer Books.
- Shintel, Hadas, Howard C. Nusbaum, and Arika Okrent
2006 Analog acoustic expression in speech communication. *Journal of Memory and Language* 55: 167–77.
- Tomasello, Michael
1999 *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.
- Turner, Mark
1990 Aspects of the invariance hypothesis. *Cognitive Linguistics* 1: 247–55.
- Wittgenstein, Ludwig
1980 *Culture and Value*. Georg Henrik von Wright and Heikki Nyman (eds.), Peter Winch (trans.). Chicago: University of Chicago.
2001 *Philosophical Investigations*. 3rd ed. Gertrude Elizabeth Margaret Anscombe (trans.). Oxford: Blackwell.
- Zbikowski, Lawrence M.
1999 The blossoms of “Trockne Blumen”: Music and text in the early nineteenth century. *Music Analysis* 18: 307–45.

- 2002 *Conceptualizing Music: Cognitive Structure, Theory, and Analysis*. (AMS Studies in Music.) New York: Oxford University Press.
- 2002–3 Music theory, multimedia, and the construction of meaning. *Intégral* 16–17: 251–68.
- 2008 Metaphor and music. In *The Cambridge Handbook of Metaphor and Thought*, Raymond W. Gibbs, Jr. (ed.), 502–24. Cambridge: Cambridge University Press.