

LAWRENCE M. ZBIKOWSKI

## Reflections on Words and Music

That there should be a close alliance between words and music is, in certain respects, hardly surprising. Both find primal expression through the human voice, and both offer a means through which humans can convey their most sophisticated thoughts. And yet each wants to go its own way: words toward the definite and concrete, and music toward the allusive and ephemeral.

In the paper I presented at the meeting of the Words and Music Association Forum at Århus University (a version of which was later published in *Signata*<sup>1</sup>) I endeavored to explain this divergence by making recourse to my recent research, which has explored the foundations of musical grammar.<sup>2</sup> That exploration took inspiration from work over the past few decades by cognitive linguists on what are called construction grammars. In a construction grammar, the basic elements of grammar consist of stored pairings of form and function, called “constructions.” If, for example, you want to ask a question in a given language you need to put words into a particular form, which we might call the “question form.” The function achieved by putting words into the “question form” is to elicit information from another person. As a whole, then, the grammatical unit known as a question puts words into a particular form (the “question form,” that aspect of the construction that corresponds most closely with what has been traditionally called syntax) to achieve a particular function (eliciting information from another person, which we can think of as representing the semantic dimension of the construction). I should want to note that this is an overly simplistic example—questions are complex structures that can take a variety of different forms, and that realize a wide range of functions—but it gives a flavor for the idea of a construction grammar. It also helps to illustrate an important difference between language and music: while it seems quite apparent that language is a useful tool for eliciting information from another person, music is not. This suggests that language and music have different functions in human cultures. If one conceives of grammatical units as combining form and function, then, it seems quite apparent that linguistic grammar and musical grammar would be substantially different from one another. From this perspective, what is remarkable is not so much the divergence of words and music but the close relationship often found between two modes of expression with quite different grammatical forms.

If it is the case that language and music have different functions in human cultures, how might we characterize these functions? In considering this question I have

1 Lawrence M. Zbikowski, “Words, Music, and Meaning,” in “Sémiotique de la musique / Music and Meaning,” ed. Per Aage Brandt and José Roberto do Carmo, Jr., special issue, *Signata* 6 (2015).

2 Lawrence M. Zbikowski, *Foundations of Musical Grammar* (New York: Oxford University Press, in press).

found it useful to draw on the work of the developmental psychologist Michael Tomasello, whose research over the past three decades has explored how children acquire language and how other primates do not.<sup>3</sup> Following Tomasello, I have adopted the position that the basic function of language within human culture is to direct the attention of another person to objects or concepts within a shared referential frame, and thus to make it possible for humans to engage in a kind of communication that facilitates cooperation within social contexts.<sup>4</sup> Music, by contrast, is not very good at directing the attention of another person to matters of shared interest—what tune would I play to let you know that you were blocking my view and making it hard for me to read?—but it is unparalleled as a resource through which various dynamic processes that are important within human culture can be represented through patterned sound. Chief among these dynamic processes are those associated with expressive gestures, with the emotions, and with the patterned movements of dance.<sup>5</sup> According to the perspective that I have developed in my recent work on musical grammar, then, one of the basic functions of music within human cultures is to provide representations for dynamic processes such as these by means of sound sequences that serve as analogs for those processes.

This perspective is framed by two further considerations. First, using a sequence of sounds to make reference to some dynamic process (which might itself not have any sonic component) requires a robust capacity for analogical thought. On the best evidence we currently have, humans are the only species to have such a capacity.<sup>6</sup> Second, human communication—whether it is through language, music, or some other medium—relies on an infrastructure for collaborative interactions.<sup>7</sup> To take a simple example, consider a situation in which I meet you prior to a lecture and point to an empty seat in the lecture hall. My gesture could be interpreted in any number of ways: I could be indicating where I intend to sit, where you should sit, where a mutual friend will be sitting, or the fact that the lecture is a popular one and only one empty seat remains. At least two things are required to interpret my gesture: a shared frame of reference (relative to which the information I provide through my gesture can be understood to apply to myself, to you, to our friend, or to the popularity of the lecture) and the knowledge that I offer this information in order to be helpful. Within the context

3 See in particular Michael Tomasello, *Constructing a Language: A Usage-Based Theory of Language Acquisition* (Cambridge, MA: Harvard University Press, 2003); and *Origins of Human Communication* (Cambridge, MA: MIT Press, 2008).

4 Michael Tomasello, *The Cultural Origins of Human Cognition* (Cambridge, MA: Harvard University Press, 1999), chapter 5.

5 In recent work I have explored connections between sequences of musical events and each of these dynamic processes. For representative work see Lawrence M. Zbikowski, "Musical Gesture and Musical Grammar: A Cognitive Approach," in *New Perspectives on Music and Gesture*, ed. Anthony Gritten and Elaine King (Farnham: Ashgate, 2011); "Music, Emotion, Analysis," *Music Analysis* 29, no. 1–3 (March–October 2010); and "Music, Dance, and Meaning in the Early Nineteenth Century," *Journal of Musicological Research* 31, no. 2–3 (2012).

6 Josep Call and Michael Tomasello, "Reasoning and Thinking in Nonhuman Primates," in *The Cambridge Handbook of Thinking and Reasoning*, ed. Keith Holyoak and Robert G. Morrison (Cambridge: Cambridge University Press, 2005).

7 Tomasello, *Origins of Human Communication*, chapter 1.

of interactions between humans, then, the communication engendered by my pointing to the empty seat involves a collaborative effort between the two of us that takes advantage of a referential frame that we hold in common, along with shared ideas about how the gesture of pointing is intended to be helpful. If this infrastructure is in place, the act of pointing (or, for that matter, uttering some string of linguistic sounds or offering an evocative passage on the piano) can serve a communicative purpose. If this infrastructure is not in place—that is, if there is no shared frame of reference, and no sense that there is an intent to communicate—the gesture (or sequence of linguistic sounds, or musical passage) will not realize its communicative function.

In that both language and music rely on this same infrastructure for communication, and both can potentially be enacted through the voice, an alliance between the two is, indeed, hardly surprising. Such an alliance should not, however, obscure the marked differences that obtain between language and music, nor justify a priori assumptions about which better expresses human thought. As a way to illustrate these differences, let us take a look at the beginning of Yip Harburg and Harold Arlen's "Over the Rainbow," written for MGM's *The Wizard of Oz* (which premiered in 1939), a work I discussed in my original paper. The opening eight bars of the song are given in Example 1. The first words of Harburg's lyrics are what the linguist Gilles Fauconnier would call a "space builder," as they establish a mental space ("Somewhere over the rainbow, way up high") within which the small story told by the song will unfold.<sup>8</sup> The words that follow begin to furnish this space, introducing an autonomous geographical construct (a "land") associated with the experiences of childhood ("I heard of once in a lullaby"). The music that sets and accompanies these words would seem to offer us nothing so definite as this collection of concepts (especially if one simply hums the melody, without the words), and yet there is much that is of substance in these opening bars. Consider Arlen's rather famous opening interval, the octave leap that sets "Some-where": this serves as a sort of musical space builder, one which opens up a registral span (from G4 to G5), a rhythmic frame (since the articulation of each of these pitches correlates with the main beats of the bar), and which suggests that G will be an important referential pitch for what follows. The octave leap also requires a kind of vocal athleticism (just *try* to sing that opening interval without taking a good breath beforehand) which is in tension with the childlike cast of the words.<sup>9</sup> Over the course of bars 3–7 the melody traces a gradual descent from the upper limit established by G5 (from E5 in bar 3, to D5 in bar 4, to C5 in bar 5, to B4 in bar 6, to A4 in bar 7, all of which are indicated by arrows on Example 1). Although stepwise motion relieved by small leaps predominates, vestiges of musical athleticism remain: not only are there leaps from G4 to E5 in bar 3 and from E4 to C5 in bar 5, these leaps also mark off two-bar units, temporal spans long enough to suggest that the whole of the melody is being carefully molded by a voice that is thoughtful and assured. If

8 Gilles Fauconnier, *Mental Spaces: Aspects of Meaning Construction in Natural Language*, 2nd ed. (Cambridge: Cambridge University Press, 1994).

9 I discuss the tension between the words and music of "Over the Rainbow" in more detail in "Words, Music, and Meaning," 157–60.

we think of this sequence of sounds as serving as a sonic analog for a dynamic process, that process would not easily correlate with the actions of a tentative spirit lost in childhood reflections; it would, however, correlate quite strongly with the energetic demeanor of a brave and confident agent ready to venture forth to discover the promised land that lies just “over the rainbow.”

G    Em<sup>7</sup>    Bm<sup>7</sup>            G<sup>7</sup>    Cmaj<sup>7</sup>    C<sup>7</sup>            Bm<sup>7</sup>    Bb<sup>7</sup>

Some - where    o - ver the rain - bow way    up    high,

5    Am<sup>7</sup>    Cm<sup>6</sup>    G/D            E<sup>7</sup>(b<sup>9</sup>)    A<sup>7</sup>            D<sup>13</sup>    G

there's    a    land that I heard of    once in a lull - a - by.

Example 1: Harold Arlen and Yip Harburg's "Over the Rainbow" (© 1939, renewed 1966, Metro-Goldwyn-Mayer Inc.), bars 1–8.

As this brief example suggests, words have the potential to prompt the construction of rich mental spaces furnished with objects and events that we can use to build a narrative. Harburg's lyrics, by themselves, introduce a mythical land in the far beyond of a sort that would enchant a young child on the edge of sleep. Music, by contrast, gives us nothing as concrete as this, but instead offers a sonic analog for a detailed and multi-dimensional dynamic process (especially if we take into account the effect of melody, harmony, and rhythm working together). Arlen's music does not fill out the picture of Harburg's mythical land as much as it suggests what it might feel like to have a deep yearning for the promise of such a land as well as a profound confidence that this land is, in some way, within reach.

As might be expected, the account of musical grammar I have developed in my recent work is rather involved, and offers a novel view of the role of music in human cultures. That said, there are any number of interesting intersections between the perspective I have developed and other work presented at the Århus University meeting, and in the following I would like to explore connections between my work and that of other contributors to this volume.

### *Sound, technology, and the infrastructure for communication*

Anette Vandsø's "Technology, Sound and Subject" takes as its main example Gilberto Zorio's *Microfoni*, in which a soundscape is created through interactions between people within a gallery space and five suspended microphones (interactions that are

encouraged through plinths placed underneath each microphone to facilitate access to the microphones). This example leads her toward a thoughtful investigation of a range of works of sound art, all of which use technology to a greater or lesser degree. Vandsø's emphasis is, quite rightly, on the way technology writ large provides a means to communicate our thoughts and ideas, an instrumentality brought to the fore by sound art (which, in most cases, would not exist without sound (re)producing technologies). Technology thus provides a means of constructing a subject, understood as the person who uses or interacts with the technology. Equally striking, however, is the infrastructure for communication that is assumed by sound art (and, to a certain extent, by sound (re)producing technologies). Zorio's *Microfoni*, for instance, relies on the shared space created by the gallery, microphones, and amplification system, all of which establish a framework for interactions with Zorio (via his installation), with other individuals in the gallery, and even with those we might imagine joining us in the gallery—that is, an infrastructure for communication. And it is worth pointing out that the interactions facilitated by this infrastructure are typically understood within an intentional context: if, unbeknownst to them, the conversation of two individuals walking through the gallery were picked up by one of the microphones we might understand both that they did not intend us to hear their conversation *and* that Zorio intended we hear the conversation as part of the shared space created by his installation. To the extent that technology facilitates the construction of a subject, then, that construction inevitably happens within a social context that is the rationale both for the technology and for attempts at communication.

### *Words and music as modes of expression*

As hinted at by my analysis of the opening of "Over the Rainbow," words and music have different functions in human cultures and realize these functions through organizing the materials of communication in different ways. The relationship between the two is nicely captured in François Staring's "Audiences in Literature," which explores the way music prompts imaginative responses from everyday listeners, responses that are expressed through words. Staring's aim is to offer an account of the listener as an engaged agent who, to a certain extent, co-creates the artwork initiated by the composer and enacted by the performer. This co-creation reflects the background and skills of the listener, as well as the context for the listening experience, and is documented (so to speak) by rich accounts of listening experiences provided by authors like E.M. Forster and Hélène Grimaud. What is particularly striking about the literary descriptions of passages from Beethoven's *Fifth Symphony* quoted by Staring is their reliance on dynamic images: in *Howards End*, Helen Schlegel hears the movements and actions of a malign goblin in the third movement of the symphony; in *Variations sauvages*, Hélène Grimaud recollects the way Beethoven's music pushed and pulled her, almost as though it were a physical force. An argument could be made, then, that the activation of these images by those reading *Howards End* or *Variations sauvages* relies to some extent on the reader's own encounters with musical evocations of dynamic processes. In the absence

of such encounters, a reader might well be able to develop some idea of how Helen's goblin moves and acts by drawing on knowledge about malevolent beings but will not be able to summon the rich dynamic image activated by Beethoven's music.

The relationship between words and music adumbrated by Staring is, to an extent, filled out in Beate Schirmmacher's "Musical Performance and Performative Discourse in Elfriede Jelinek's *Die Klavierspielerin*," which shows how Jelinek's characterization of her troubled protagonist is informed by the embodied realities of musical practice. To be sure, some of the ideas activated by Jelinek's prose that Schirmmacher explores might be associated with a range of human endeavors. To play a musical instrument well involves a kind of athleticism—albeit one typically focused on small, rather than large, muscle groups—which is also evident in actual athletic activity. Thus the disciplined body of the musical performer is also the disciplined body of the gymnast or the long-distance runner, and as such all could have served as the basis for Jelinek's metaphors. Where musical practice is different (and where it is similar to the practice of dance or the theater) is in the assumption that bodily discipline is in the service of a species of communication. And yet it is communication—at least about that which matters most—that is beyond the protagonist of *Die Klavierspielerin*. Thus music, representing both the ineffable (in the sense of "that which is beyond words") and the embodied (through the physical act of practicing for performance and through the representation of dynamic processes), comes to be a wraith that haunts the novel.

It is the possibility of pinning down this elusive wraith through graphical means that Sybille Krämer explores in her "Script and Sound." Krämer notes how the technology of writing takes time-dependent media *out* of time, and facilitates systematic accounts of those features that are captured within the two-dimensional frame of the printed or written page. Although she draws her example of this practice in a musical context from diagrams that appear in René Descartes's *Compendium Musicae* of 1618, one can also find such diagrams in works as early as the Euclidian *Sectio Canonis*, which dates from around the third century BCE.<sup>10</sup> On the one hand, diagrams that graphically represent musical relationships are clearly part of a process of systematization; on the other hand, they also serve as what Edwin Hutchins has called material anchors.<sup>11</sup> The purpose of material anchors—which are widespread across human cultural practices—is to provide a stable material referent for often ephemeral or complex concepts; they also help to guide interactions between humans negotiating those concepts.<sup>12</sup>

10 A translation of the *Sectio Canonis*, with commentary, is provided in Andrew Barker, ed., *Greek Musical Writings*, vol. 2: *Harmonic and Acoustic Theory* (Cambridge: Cambridge University Press, 1989), 190–208. As Barker observes (194, n. 9) the manuscript tradition for the diagrams is not as consistent as for the manuscript proper; it does seem, however, that these diagrams were incorporated into the manuscript at a relatively early date.

11 Edwin Hutchins, "Material Anchors for Conceptual Blends," *Journal of Pragmatics* 37, no. 10 (October 2005): 1555–77.

12 Edwin Hutchins, "The Distributed Cognition Perspective on Human Interaction," in *The Roots of Human Sociality: Culture, Cognition and Interaction*, ed. N. J. Enfield and Stephen C. Levinson (Oxford: Berg, 2006), 375–98. For an extended argument on the importance of material objects for human thought across archeological history, see Lambros Malafouris, *How Things Shape the Mind: A Theory of Material Engagement* (Cambridge, MA: MIT Press, 2013).



Thus pieces of metal or even paper can be used to anchor the rather intangible notion of “money,” or dots amidst lines to represent musical pitches. The potential of material anchors to aid thought is in fact mentioned by Krämer, who observes that graphism, spatialization, and visualization all have the potential to stimulate the creative production of music. What is perhaps most striking, however, is the way such anchors can even threaten to displace that for which they stand: as shown by Lydia Goehr, the “work concept” that developed around European musical practices during the late eighteenth and early nineteenth centuries owed much to the possibility of capturing the salient features of musical utterances through graphic means—that is, through the score.<sup>13</sup> Although, as Krämer notes, the distortion of the temporally-dependent nature of media introduced by graphism applies equally to words and to music, an argument could be made that this distortion is more pronounced in the case of music. Words, after all, give humans a means to transcend the boundaries of time through allowing us to construct notions of a “past” and a “future.”<sup>14</sup> Music has nothing to equal this, but through the interactions of resources like melody, harmony, and rhythm, makes possible the creation of a present of unparalleled richness.

### *Sonic resources in music and language*

Another aspect of alliances (and misalliances) between words and music considered by a number of contributors to the conference concerned the resources that each medium deployed in the service of communication. Birgitte Stougaard Pedersen, in her “The Generalpause and the Enjambment,” explores a sort of “negative” resource exploited both by music and poetry: silence. If we conceive of expressive media as exploiting figure-ground relationships (with the “message” of the medium a figure situated against some ground), it seems clear that both music and words rely, to some extent, on silence as a contrasting element against which to array their sounds (something particularly evident if we try to attend to either music or words in a noisy environment). Following the composer and theorist Edward T. Cone, we might even conceive of silence as a kind of frame that sets a musical composition off from its surroundings, just as a picture frame sets a picture off from the wall on which it hangs.<sup>15</sup> The silence with which Stougaard Pedersen is concerned, however, is internal to a work (be it of music or poetry), and typified by the *Generalpause* and poetic enjambment. On the one hand, such a silence provides a means to interrupt the flow of musical events or the metrical structure of poetic verse, signalling a break in the dynamic process summoned by the preceding events. On the other hand, a case could also be made that silence of this sort is not unlike the watercolorist’s “white.” Given the nature of watercolors, the only practical way to represent a white color (so that one can depict, for instance, the sail of a sailboat) is to apply no color to the surface upon

13 Lydia Goehr, *The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music* (Oxford: Clarendon Press, 1992).

14 Tomasello, *Origins of Human Communication*, 284–88.

15 Edward T. Cone, *Musical Form and Musical Performance* (New York: Norton, 1968), 16–17.

which one is painting—the visual equivalent, then, of the musician’s or the poet’s silence. In her 2007 study of the functions of silence in music Elizabeth Margulis offers other ways to think about how silence is deployed in music (and, by extension, in poetry). Margulis notes that, in addition to serving as a boundary or an interruption, silence can also reveal the operation of the inner ear (that is, the “hearing along” that is the mark of close attention to a sequence of sounds), promote attention to the listening process itself, and (perhaps most strikingly) “communicate” through imitating the kinds of silences that occur in expressive speech.<sup>16</sup> As Stougaard Pedersen suggests, the resources of silence play a role both in music and in poetry, and reveal much about the way sound—against the backdrop of silence—shapes our understanding. Molly McDolan is equally concerned with silence, but here as a multidimensional concept rather than simply as the absence of sound. In her “*Stille* in the Lutheran Baroque” McDolan explores various means by which Baroque composers represented the notion of *Stille* which, within ecclesiastical texts, was associated not only with silence but also with the context for seeking a quietude of the spirit. As such, composers summoned *Stille* through coordinating a variety of musical resources, including “rests or a pedal tone in the bass line, a sustained note in a low register in the solo voice, a generally reduced texture, and active harmonically-stable figuration in the accompaniment,” all of which Johann Sebastian Bach uses in his setting of *Stille* in his cantata *Gott, man lobet dich in der Stille* (BWV 120). Perhaps of equal importance, Bach juxtaposes these compositional strategies with those he uses to set the word *lobet* (“praise”) immediately prior to *Stille*, which are nearly their direct opposite: a moving bass line, an incredibly florid melody, a rich orchestral texture, and harmonies that suggest a directed, progressive process. McDolan proposes that this approach to the text exemplifies what she calls “musical iconography,” through which music is used to summon visual images. This notion fits quite well with my idea that one of the basic functions of music is to provide sonic analogs for dynamic process—in the case of Bach’s cantata, I would see these processes as beginning with energetic praise that then fills the stillness of quiet, prayerful reflection with something like an electrostatic charge—and is further supported by the extensive corpus analysis that informs McDolan’s study.

Although Axel Englund does not focus directly on the sonic resources of music and language in his “Of Great Pitch and Moment,” his analysis of the importance of “moment” in opera studies could be seen to highlight the contribution music makes to theatrical works. As Englund observes, a complex discourse has built up around the notion of “moments”—typified by an absorbed engagement equal to, if not exceeding, sexual climax—which tends to direct attention away from the multileveled and at times overlapping processes that create the possibility of such climactic experiences. What we need, Englund proposes, is a hermeneutics of opera that allows meaning to be contingent and changeable. From the perspective developed in my work, I would suggest that the tendency against which Englund reacts, which has its consequence in

16 Elizabeth Hellmuth Margulis, “Moved by Nothing: Listening to Musical Silence,” *Journal of Music Theory* 51, no. 2 (Fall 2007): 245–76.



assertions of meaning inadequate to the aspirations of opera, is in part a result of trying to reduce all kinds of meaning to that which is concretized in language, and which resists the ongoing dynamic processes that are central to musical expression.

*Reflections on music and words*

As was evident in the discussion of linguistic and musical construction grammars with which I began these reflections, the perspective that informs my research is that of a music scholar. My contribution to scholarship, then, has been to give an account of the unique expressive resources music has provided to human cultures—a uniqueness evident in the fact that every human culture of which we have knowledge has both music and language. In consequence, the comments I have made on these provocative contributions to the meeting of the Words and Music Association Forum at Århus University have focused more on *music* than on *words* (although I believe the contrast between the communicative resources each offers can be instructive for our study of both). Needless to say, scholars whose work focuses more on words than on music will almost certainly take different things away from these papers, as they do from their study of words and music. That said, I believe that continuing to think about what music brings to human life—with both “music” and “life” understood in all their complexity—can help us better understand what words can and cannot accomplish, and thus to enrich our study of music and words, as well as words and music.