



Schemas, grounds, meaning: On the emergence of musical concepts through conceptual blending

Musicae Scientiae
2018, Vol. 22(1) 57–71
© The Author(s) 2017
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1029864917711218
journals.sagepub.com/home/msx


Mihailo Antović

University of Niš, Serbia

Abstract

This article offers a new theoretical approach to the conceptualization of music, based on Conceptual Blending Theory, with a reinforced role ascribed to the constructs of generic space and the grounding box. Three analyses of typical conceptualizations of music from prior experiments with children and adults are provided to postulate that the ultimate linguistically reported concept comes from blending the intramusical Gestalt (input space 1) with a rich image from an appropriate experiential domain (input space 2). However, the mapping is not haphazard, but rather based on the invariant structure in the generic space, which takes the form of an image-schema family. In the blended space new conceptual elements emerge: one such typical resultant concept generates the idea that music “moves”, and in specifically articulated ways. While more basic verbal reports from experiments may be constrained by image-schema families alone, richer descriptions additionally require the theoretical notion of the grounding box, which hosts experiential information that participants add to the description as they progress in building musical meaning. The proposed model relativizes two common dichotomies in music cognition: (1) the distinction between “intramusical” and “extramusical” meaning, since both participate in the process of creating the ultimate blended concept; and (2) the strict divide between universalism and linguistic relativity in musical concept formation, since the present proposal has sufficient theoretical constructs to account for *both* schematic invariants and experiential diversity.

Keywords

conceptual blending, conceptualization, generic space, grounding box, image schema, music

Conceptualization is one of the major questions of cognitive science. Philosophers, psychologists and linguists, among others, have long wondered where concepts come from, what they are grounded in, and whether they share “building blocks” cross-culturally, even if they seem different in various languages. Proposed answers have embraced a variety of positions, from universalism to linguistic relativity, from nativism to neo-behaviorism to constructivism. The present contribution aims to suggest that research of musical concepts can be an excellent vehicle for addressing some of the big dilemmas of meaning construction in general. More particularly, this

Corresponding author:

Mihailo Antović, Faculty of Philosophy, Center for Cognitive Sciences, University of Niš, Ćirila i Metodija 2, 18000 Niš, Serbia.

Email: mantovic@gmail.com

work proposes to use a model based on the Conceptual Blending Theory (CBT), with reinforced roles ascribed to the constructs of *generic space* (Fauconnier & Turner, 2002) and *grounding box* (Coulson & Oakley, 2005). In an attempt to bridge the gap between disparate epistemological approaches mentioned above, in the present proposal the blueprints of a system are provided that might capture *both* cross-cultural diversity and schematic invariants underlying some frequent, yet seemingly different musical concepts across languages and populations. Beyond that, three groups of common linguistic descriptions of short musical fragments will be reanalyzed, as obtained in experiments from naive participants of various age, cultural and linguistic background, and cognitive status (Antović, 2009; Antović, Bennett, & Turner, 2013; Antović, Mitić & Benecasa, 2017; Antović, Stamenković, & Figar, 2016). Including this introduction, the article consists of four parts: section two provides essential information on conceptual blending, along with its recent applications to music cognition, and then revisits some of the most interesting verbal labels that participants used to describe musical concepts in previous studies. Section three suggests why commonly employed theoretical approaches, such as that of the Conceptual Metaphor Theory, stop halfway in explaining the experimental data. It then provides three analyses to argue that a system based on conceptual blending, with an enhanced role assigned to the generic space and grounding box, can more convincingly account for the motivation behind the conceptual variability which regularly emerges in experimental work. Section four offers directions for future research and introduces the ongoing experiments by Antović and associates, which use the approach proposed here not merely as a means for post-hoc analysis, but also as a tool for postulating empirical hypotheses.

Blending and musical conceptualization: Theoretical background and prior work

Conceptual blending (CB; Fauconnier & Turner, 2002) focuses on operations at high levels of cognitive integration, and addresses the construction of meaning in various cognitive domains – including linguistic metaphors, counterfactuals and puns, as well as the aesthetic appreciation of literary images. Essentially, it proposes two sets of online conceptual packets, *input mental spaces*, whose elements selectively *map* onto one another, where novel, *emergent* structures occur in the newly-created, *blended* space. In addition, there is commonly a *generic space*, which hosts preconceptual, often spatial, topologies common to the two inputs, holding the system together and motivating the mappings in the first place. The Sphinx, the snake from the Garden of Eden, “space-time” in physics, “consciousness as the tip of the iceberg” in Freudian psychology, or unconventional expressions such as “That surgeon is a real butcher”, are but a few structures embodying the process of blending. In this last example, mapping the semantic elements of the “surgeon mental space” (a conscious professional using a sharp instrument to cut human flesh in order to heal) onto those of the “butcher mental space” (an equally conscious professional using a sharp instrument to cut animal flesh in order to carve and ultimately eat meat) results in a sense of imbalance in the blend, since even the most technically proficient butcher cannot use a cleaver to heal. This discrepancy between *means and ends* in the input spaces engenders the emergent semantic element of *incongruency*, which is crucially responsible for the interpretive power and emotional appeal of this unusual metaphorical expression (see Grady, Oakley, & Coulson, 1999, or, for an alternative analysis, L. Brandt & P. A. Brandt, 2005).

Blending has excelled as a functional approach to linguistic semantics, but it has lately been applied more broadly in cognitive science, including music cognition. In terms of semantic phenomena which musicologists often label *extramusical*, i.e. relating musical structure and the world of experience (cf. Koelsch, 2012), Zbikowski (2002) has discussed text painting, in which

a musical and a linguistic structure, such as a trill and the word “trembling”, blend to create an emergent, augmented, perhaps even bodily effect of trepidation in the audience. Accordingly, in more recent work Zbikowski has increasingly argued for the *bidirectionality* of musical-textual mappings (Zbikowski, this issue). Cook (2001), Sayrs (2003), Chattah (2006), Tasoudis and Vouvaris (2016, and this issue), and Antović (in press) have used blending to discuss music semiotics, typically interpreting extramusical meaning in program music. These contributions used blending to analyze, respectively, the impact of television commercials, the connection of lyrics and music inspired by a well-known literary novella, and the semiotic effects of film music. Recently, the connection between blending and musical emotions has also been studied, providing a fresh theoretical perspective on one of the most widely debated questions in the psychology of music in general (Spitzer, this issue). On the other hand, P. A. Brandt (2008), A. Brandt (Brandt & Eagleman, 2017), and Antović (2014) have also analyzed intramusical constructs, such as the integration of rhythmic and melodic patterns, counterpoint, pitch hierarchies, or complex meter as instances of blending. From an equally intramusical perspective, recent research has seen excellent progress in the computational corroboration of the construct of blending, in music mostly in terms of chord sequencing and harmonization (Cambouropoulos, Kaliakatsos-Papakostas, & Tsougras, 2015; Eppe et al., 2015; Zacharakis, Kaliakatsos-Papakostas, & Cambouropoulos, 2015); One of the most remarkable achievements of this program is the fact that an AI routine has managed to blend two well-formed cadences from the classical repertoire and infer the “emergent” tritone substitution, simulating the actual historical development of this musical concept – from classical to jazz harmony (Zacharakis, Kaliakatsos-Papakostas, Tsougras, & Cambouropoulos, this issue). The applicability of the intra-/extramusical distinction has been critically questioned from a CBT perspective (Stefanou & Cambouropoulos, 2015), where both types of phenomena have been successfully analyzed in the same musical composition (Tsougras & Stefanou, 2015, and this issue). Finally, the importance of conceptual blending for *performance*, rather than just perception or conceptualization has been discussed in relation to so-called “intermedial blends” (Stefanou, this issue).

All these studies ask important questions about musical conceptualization, for instance, on any qualitative difference between intramusical concepts, such as a five-note theme taken as an absolute piece of music, and extramusical ones, e.g. the same theme being related to an event, character, or sensation from the world of experience, often expressible in language. For musicologists who base their approach on Cognitive Grammar – a school in linguistics denying the sharp divide between form and meaning – this distinction collapses (Zbikowski, 2002). For others, it remains theoretically important (Kühl, 2008, p. 86). Some related earlier work has also taken a traditional metalinguistic perspective and insisted on the distinction (Antović, 2014), yet here more recent experimental data are considered, in search of a theoretical approach that might indeed render the dichotomy unnecessary. Partly independently of the intra-/extramusical divide, the question of universals remains equally important. When musical concepts are constructed in different cultural circumstances, as in a scale that goes “up and down” in one language but becomes “thinner and thicker” in another – one might ask if the conceptualization is chiefly motivated by the speaker’s knowledge of the mother tongue (Dolscheid, Shayan, Majid, & Casasanto, 2013) or if there might be higher-order, schematic invariants beneath apparently disparate cross-linguistic conceptual choices (Antović et al., 2017, partly based on Jackendoff, 1990).

A series of studies with Serbian, Roma, and American sighted and blind children, and musician and nonmusician students, has revealed significant cross-linguistic differences in the conceptualization of basic musical elements, such as pitch distances, movement, dynamics, and (dis)harmonious chord sequencing. In two such studies, 10-year-old naive participants were

exposed to diametrically opposed musical stimuli, such as an ascending and descending musical scale, and asked to verbally describe the difference between the two, as best they could. The raw responses were then classified into higher-order, metaphorical categories. Upon the analysis, some interesting differences emerged among the preferred conceptualizations in different groups. For instance, for Serbian children, two tones an octave apart were typically “high and low”, while for Romani pupils they were “big and small” or “thick and thin” (Antović, 2009). The former group viewed pitch movement predominantly as going “up and down”, the latter as getting “bigger and smaller” or going “to the goal and back”. In a replicated study with non-sighted children, musical movement was occasionally defined as a change from “heavier toward lighter” or even occurring “in a circle” (Antović et al., 2013). Similar results emerged from descriptions of other musical constructs: a staccato and a legato line was commonly “abrupt and linked” or “hopping and walking”; a piano and forte note sounded “weak and strong” or “letting go and pushing”. Finally, in the most recent studies by the same group, which asked nonmusician students to provide descriptions of any meaning provoked in them by pieces of programmatic music, with or without prior linguistic prompts, an actual musical excerpt (*So grüss ich die Burg* from *The Ring of the Nibelung*) was found to resemble “growing tension” resulting in a “path of two armies that are about to clash”, but also reminded some of “growth, since the tonality elevates” (Antović, 2016; Antović et al., 2016).

This diversity of verbal reports leaves two choices for interpretation: more directly descriptive, in which one highlights the obvious surface differences to support the thesis of linguistic relativity; or more theoretically speculative, where one performs a post-hoc linguistic analysis in search of potential deeper semantic constraints – ideally revealing some invariants beneath the different verbalizations. The dilemma of course does not have consequences only for musicological debates, but may be fundamentally important for broader discussions in semantics and cognitive science. In linguistics, for instance, it has defined entire paradigms. The breakup between Chomskyan Generative Grammar (Chomsky, 1981) and Lakoffian Cognitive Linguistics (Lakoff, 1987) centered largely on the problem of semantic atomism and universals, and remains a matter of fierce debate even today (Pinker, 2013 vs. Evans, 2014).

The remainder of the paper attempts to suggest that in (musical) conceptualization studies “intra-/extramusical” and “universalist/relativist” should not be viewed as dichotomous poles. Rather, a conceptual blending approach may be useful in capturing both shared schematic invariants and experiential cross-linguistic differences in (musical) concept formation. To take the conceptualization of scales as an example, the individual verbal descriptions have certainly differed in previous studies: the music indeed “went up and down”, “thinned and thickened”, “became smaller and bigger”, “moved from heavier to lighter” or “ran in a circle”. Yet instead of taking these descriptions as a clear sign of linguistic relativity, one can also view them as a motivation to look for higher-order constraints. In other words, the actual (different) verbal descriptions do not point only to what *is* possible, but also to what, in principle, *is not*. For instance, pitches can hardly be naturally compared to “apples and bananas” (Zbikowski, 2002, p. 70). Thus, focus on the nature of conceptual constraints is as legitimate a research question as is descriptive cross-cultural comparison. In the scales example, whether the tones “move” from a low towards a high position in space, “run” in a circle or along a vertical line, “shrink” horizontally or by all axes, or finally stepwise “let up” physical pressure, there seem to be at least three schematic notions underlying the apparently diversified conceptualizations – FORCE, DISCRETE DISTANCE, and PATH (Antović, 2014). However the particular concept may be finally framed – and this is certainly culture-specific and has a lot to do with the participant’s native language and personal experience – all five conceptualizations seem to inhere these proposed notions. After all, for the static tones to start “moving” – which is a conceptual, and not an auditory

phenomenon – some (mentally represented) force needs to be applied, after which the representations of the tones metaphorically “traverse” discrete distances. In turn, these sequential, adjacent distances result in a path trajectory, again irrespective of whether the conceptualized movement is linear (horizontal, vertical, diagonal), circular, or more abstractly based on the succession of pressures.

These proposed schematic constraints mirror one of the fundamental postulates of Cognitive Linguistics – *image schemas*, pre-conceptual Gestalten providing the basis for structuring abstract concepts (Hampe, 2005; Johnson, 1987). Of course, linguists typically look for schemas in verbal utterances. For instance, we can locate *FORCE* and *PATH* in many expressions in language which have literally nothing to do with force or movement, such as “I have *strong feelings* for her” or “Our *paths* have crossed again”. Yet similar schematic elements can be found in other domains, e.g. purely visual (Antović, 2010) or mathematical (Lakoff & Núñez, 2000). Applying them to cognitive musicology is not a novelty, either. Larson (2012) has provided a detailed theory of cognitive forces in music, while the notion of *PATH* and other spatial topologies have been ubiquitous, and discussed in relation to image schematicity since the work of Saslaw (1996), Brower (2000), Zbikowski (2002), and Johnson and Larson (2003). What can be considered new in the present approach is (1) the attempt to introduce image-schematic constraints as a *common ground* beneath a number of cross-linguistic and cross-cultural possibilities for describing the same musical concepts, and (2) the methodological decision to base such an analysis on actual utterances obtained from actual participants in behavioral experiments. How the proposed system works in practice will be explained in the next section, which analyzes the most common linguistic descriptions of three musical stimuli from previous work with children and adults, as follows: (1) an isolated five-tone staccato and legato sequence played side by side as musical opposites (Antović, 2009); (2) the staccato fragment from Grieg’s *In the Hall of the Mountain King* described by naive nonmusician students as part of a semantic experiment; and (3) the violin trill excerpt from Vivaldi’s spring movement from *The Four Seasons* verbally portrayed in the same experiment (Antović et al., 2016). The goal will be to affirm blending as the theory of choice for explaining the musical conceptualization process.

Three analyses

The previous section introduced image schemas as the theoretical construct constraining the range of musical conceptualization. Yet, image schemas in themselves are not sufficient. For instance, even if the description of scales is based on a *PATH* topology, actual participants in experiments do not just talk about “traversed paths”. Typically they provide richer imagery, often personified/anthropomorphic (Watt & Ash, 1998). Hence for one to describe an upward major scale, for example, as “running in a circle” or “moving upstairs”, one needs at least to (a) *make an intramusical concept* in input space 1 (create a unified mental whole, or a Gestalt, out of the eight successive pitches) and (b) *evoke from long-term memory an extramusical concept* in input space 2 that contains elements which could appropriately map onto the intramusical concept. This may be an image of a little square revolving in a circle, of a person walking upstairs, of a physical object becoming lighter, etc. At this point the mapping starts. The major earlier approach in Cognitive Linguistics, known as Conceptual Metaphor Theory (Lakoff & Johnson, 1980) typically stops at this point and just assumes a series of connections between musical and extramusical domains. The “vertical movement” metaphor could map the music onto a series of stairs on a staircase, with the particular organization of the eight steps in the stairway standing for the musical key, and the movement of a person along the stairs giving us the succession of pitches, i.e. the sense of musical movement. In the “circular” situation there

would be no staircase, but rather a round space for the person to traverse; similarly, in a “teleological” metaphor the person would go from a more abstract “beginning” to an “end”, while in the “horizontal” metaphor there would be a flat surface along which one would walk “forward and backward” (cf. some mappings in Johnson & Larson, 2003; Antović, 2009).

But this analysis is not explanatory enough, for two reasons. First, it assumes an “intuitive” connection between the music and the experiential domain, instead of providing an explanation for the why and whence of such a connection. It does not show what it is that holds the system together, motivating the mappings in the first place. It also leaves open how the musical movement came about. While it is true that the musical input space may itself be viewed as a conceptualization, a mental construct rather than the physical succession of pitches, the sound stimulus motivating this intramusical concept still did not “move” anywhere: all eight tones were played from the same, static headphones. On the other hand, the extramusical, conceptualized “agent” in input 2 indeed changed his or her position, perhaps going upward along eight stairs. So the question is how the static tones and the dynamic experiential movement could saliently map onto one another in the ultimate concept/mental representation, in which the music itself seems to be moving.

Therefore just two mental spaces/conceptual domains are not enough for successful mapping in this case. To allocate the factors constraining the possibilities in the final concept, such as the postulated schemas of FORCE, DISCRETE DISTANCE, and PATH, a third space is required. Its purpose would be to preclude the “apples and bananas”, i.e. to “tell” the input spaces what, in principle, is possible, and what is not. On the other hand, to explain how the musical movement suddenly appeared, one needs to introduce the notion of emergence. In other words, the musical percept alone did not move anywhere. The movement was the conceptual *result* of the mapping operation and it *emerged* in the fourth, blended space.

The allocation of image schemas to one of the mental spaces in the blending network, and more generally the role of the generic space, are controversial issues in Conceptual Blending Theory (for criticism, see L. Brandt & P. A. Brandt, 2005; P. A. Brandt, 2008; Coulson & Oakley, 2005). In terms of musical concept construction, the position advocated here is in line with Hedblom, Kutz, and Neuhaus (2015) who claim that constraints on meaning generation must be image-schematic and in principle localizable in the generic space (they sometimes call it *the base ontology*). They also rightly note that image schemas motivating concepts typically come together in ordered sets, *image-schema families* (Hedblom, Kutz, & Neuhaus, 2016, partly following Mandler & Pagán Cánovas, 2014). What is added in the present proposal, however, is that image schemas alone are not always enough as the motivating force. While for more basic cases, such as the definition of “clean” music-theoretic concepts, they may be sufficient, more elaborate musical descriptions will require additional contextual grounding. This is exactly where the schematic invariants and the experiential factors meet, allowing for a good methodological opportunity to account for both more universal and more cultural aspects of concept formation in a single theory.

The first set of typical participant responses that is analyzed here is quite schematic. It comes from two previous studies with young participants, looking into image schemas underlying elementary musical concepts, such as pitch distances, successive movement in scales, or tempo variations (Antović, 2009; Antović et al., 2013). This research program found that children had given predominantly metaphorical descriptions of music, with some cross-linguistic and cross-cultural differences, yet with a possibility to infer more abstract, higher-order invariants beneath the responses. The particular example of interest here relates to the conceptualization of a five-tone sequence played on a computer sample simulating a violin timbre. The first sequence was played *staccato*, and the second *legato*, and then the participants (nonmusician

and musician Serbian, Roma, and US sighted and blind 10-year-olds) were asked to describe in brief terms “what the first (x) and what the second part (y) were like”. One conceptual class stood out as the mode, and was statistically significantly much more common than the remaining ones. More particularly, in 53.4% of the cases, the children talked about some form of interrupted and uninterrupted (human) movement to describe the sequences. The remaining classes were much less frequent, describing these pitch relations as qualities – 14.4%, unrelated extramusical descriptions – 6.7%, or sizes – 5.6%. The three most common individual conceptualizations were movement which was “interrupted and continued”, “abrupt and linked”, and “hopping and walking”. These also employ the metaphor of musical movement, arguably grounded in the same schemas responsible for motion in the scales example: (FORCE/MAGNITUDE)/PATH (Antović, 2014). However, yet another invariant seems to be present here: the LINK schema, where individual elements of music that move (such as tones) are either “disconnected” (interrupted, abrupt, and hopping) or “connected” (continued, linked, or walking) to one another. The schema is also noticeable in conventional musical notation, as staccato is usually presented with dots (punctuation) and legato with a solid curve (slur) above or below the notes. Thus, the underlying generic set of schemas for this type of musical motion could look as follows: [(FORCE/MAGNITUDE)/PATH] + {LINK}, where the optional schema of LINK serves to *modify* (articulate) the musical movement. Along with the methodology introduced above, I propose to allocate these schematic constraints to the generic space. The musical concepts (the conceptualized five-tone staccato sequence – “x” and legato sequence – “y”) will comprise the intramusical input space 1, and the three extramusical descriptions involving different types of movement would belong to the input space 2 (for simplicity of presentation, they are given below in a single blending network). The emergent property in the blend is again that of *musical movement*, which is now *articulated* in two ways – with the constituent elements either less or more interlinked (see Figure 1).

Out of the three verbalizations, “interrupted and continued movement”, just like “abrupt and linked movement” seem strongly schematically motivated: FORCE, DISTANCE and PATH schemas engender the notion of musical motion, while the antonymic adjectives used by the participants are almost synonymous with the presence or absence of links. The sense of moving music emerges, in which the motion is articulated in two dichotomous ways. However, the third description, “hopping and walking” does not exclusively reflect the generic image-schematic constraints, but rather invokes a richer image, that of *animate*, likely *human* movement. It is in such descriptions that experiential factors – individual, linguistic, cultural – become an important part of concept construction, resulting in remarkable conceptual creativity. To illustrate this, the third most common individual description of staccato and legato given by our participants was indeed “hopping and walking”, but other enriched conceptual images they produced, reflecting the same underlying relation, included “hops and walks”, “hopping and treading”, “hopping and strolling”, “running away and rushing”, and “sneaking and expectation”. Clearly, these resulting, blended concepts were no longer *only* defined by image schemas. While still based on the same set of four schematic constraints, they now incorporated additional experiential knowledge.

Such a tendency is much more conspicuous when actual examples from the musical literature are used as experimental inputs. With this in mind, a recent study (Antović et al., 2016) tested the interpretation of one of the most remarkable staccatos of the common practice period – the beginning of Grieg’s *In the Hall of the Mountain King*. Two hundred and one non-musician students, unfamiliar with Grieg’s musical program, were asked to provide a one-sentence description of the musical motive they had heard. Upon coding and calculating inter-rater agreement, it turned out there were 76.3% responses classifiable as based on image-schematic

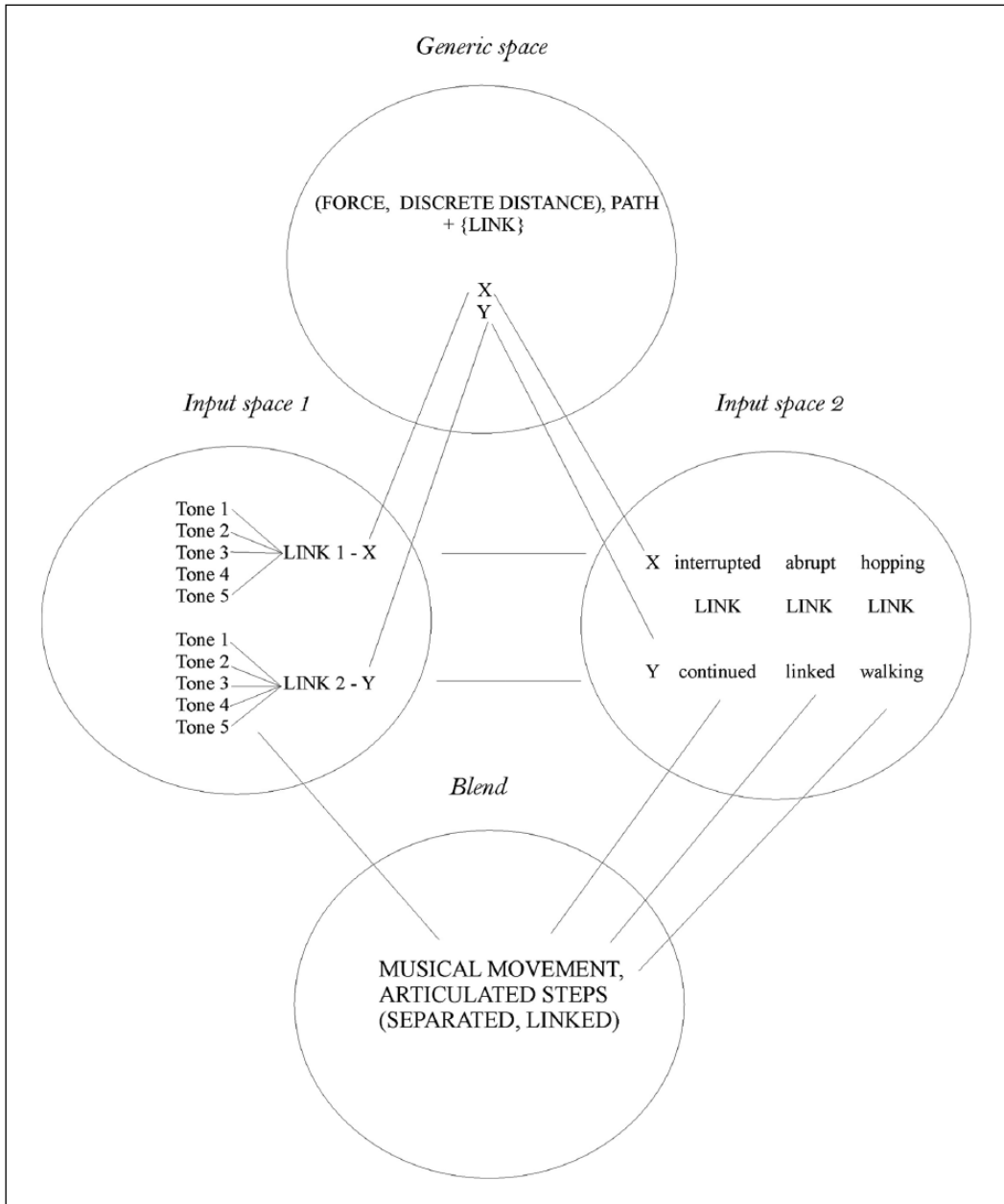


Figure 1. Generic legato/staccato conceptualization.

notions, 64.1% of which involved some form of human or animal movement, typically sneaking, stalking, walking, approaching in a concealed manner, a quest for someone, a person moving sluggishly, very comic treading of an animal, etc. Three most common conceptualizations were “heavy steps”, “sneaking”, and “stalking someone”. Again, the descriptions most likely focused on the staccato in the string section, and involved schemas underlying musical

movement (DISCRETE DISTANCE, PATH and FORCE, this last one most notably in the invocation of “heavy” steps). Appropriately, just like in example 1 above, such a schematic movement was further articulated by the LINK schema: the participants’ lexical choices typically stressed the *separated, interrupted* nature of leg movements, focusing on individual steps, sneaking and stalking but not just on unmarked motion, for instance, walking or moving around. Yet, in this set of responses one notices that in the final concept the image-schematic basis becomes less directly employed. Rather, quite a bit of additional contextual information is added to the ultimate description. Hence participants often named animal species conducting the sneaking movement or stalking their pray, described scenes in which thieves were slowly approaching a house with an intent of breaking in, pointed at the arrival of an important, mysterious character, etc. Therefore, at this point the image schemas in the generic space provided just the skeleton for the final conceptual construct, yet much richer experiential knowledge was added to the final description of the music.

If one wishes to incorporate these added elements in the consideration of musical concepts, not even the four-space blending system is sufficient. The four interconnected image schemas in the generic space may indeed constrain the possibilities for the description of the staccato. The intramusical input space straightforwardly interprets the entire 13-note theme as a single Gestalt, and the extramusical space provides a rich narrative based on the image schemas in the generic space, chosen by the participant from an almost unlimited pool of experiential options. This last tendency convincingly supports frequent claims in the literature that the range of musical semantics is potentially unlimited (Swain, 1996, p. 140), where the linguistic descriptions of music do not “*constitute* musical meaning; they *open toward it*” (Kramer, 2011, p. 15, emphasis in original). The emergent property of the music in the blend is again articulated movement, with steps clearly separated from one another. Yet, where do so many rich final images come from? The answer must be that they are *contextually based*, coming from the participant’s individual experience and cultural circumstances. This is why two prominent blending theorists, Coulson and Oakley, have proposed to add to the blending system the notion of the *grounding box*. This is not a mental space, but rather a set of contextual “assumptions that need not be explicitly represented by speakers, though they influence the way that meaning construction proceeds” (Coulson & Oakley 2005, p. 1517). The idea behind the concept of the grounding box is that in real-time meaning construction, conceptualization does not stem from the understanding of schematic or metaphorical relations alone, and not even only from blending the elements of appropriate mental spaces. Rather, to account for the complexity of the meaning generation process, a theorist should take into account the multitude of operations working in so-called “background cognition”, such as the interlocutors’ immediate contextual experience, overall cultural milieu, personal background, or values (in relation to music, the grounding box is analyzed in detail in Antović, 2016).

In terms of the experiment in focus here, the participants were just given the musical pieces to describe, without any additional context provided. They had little to start from, but the image-schematic constraints. While many have obviously adhered to these, the rest of the interpretation they needed to construct themselves, which is why they created their own contextual frames. A possible blending analysis of such responses would therefore need to introduce the grounding box to the system, to account for the contextual elements that interviewees in the experiment added to the interpretation; for example, participants, settings, and circumstances (see Figure 2).

An even more revealing set of descriptions was given with regard to example 3 – a 28-second section of violin trills from Vivaldi’s *Spring*, commonly interpreted in music criticism as an onomatopoeia for the chirping of birds. The same study with 201 student participants unfamiliar

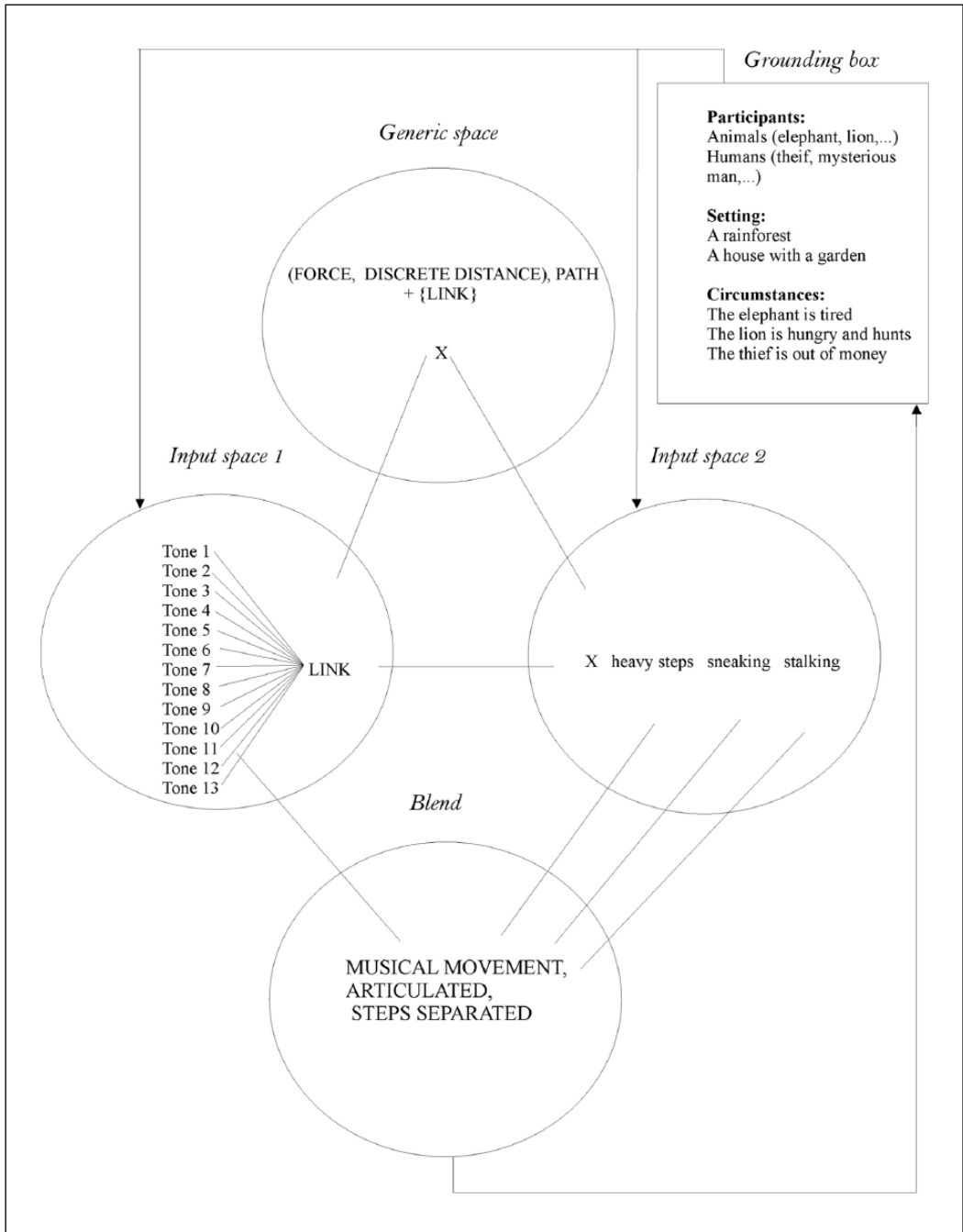


Figure 2. Conceptualization of the staccato in Grieg's *In the Hall of the Mountain King*.

with the musical program (Antović et al., 2016) revealed a striking result in which onomatopoeic responses were by far outnumbered by image-schematic ones. Presented in numbers, there were 134 image-schematic descriptions of the segment (72.43%), as opposed to 22

onomatopoeic verbalizations (11.89%). Out of the latter, only 11 had to do with birds and their chirp – the rest were unrelated (“metallic sound”, “chalk screeching”, “people chattering”, “glass shattering”, “a bee buzzing”, etc.). In terms of those responses that contained references to birds, it turned out again that onomatopoeias (11, e.g. “twitter”) were outnumbered by image-schematic notions of movement (e.g. 17 descriptions of the “flutter of wings”). This result in itself warns against the insistence on onomatopoeia in musical meaning construction and justifies an approach that would favor cognitive processes less direct than mere imitation. Three typical individual verbalizations here involved the “swarm of bees”, “flutter of wings”, and “flight path of mosquitoes”, extramusical images allocated to input space 2 in Figure 3. The intramusical input space 1 consisted of the violin trills, i.e. of musical ornaments comprising quickly alternating lower and higher tones, separated by a half or full step. In Figure 3 below, only four tones from a single trill are captured, and the notion of REPETITION is added to designate that the same pattern repeats through several measures (14, 17–18, 23–26), providing the Gestalt for the intramusical concept in space 1. The generic space would need to contain the schemas shared by the musical concept and the three extramusical descriptions. It seems that the phenomenon of musical movement can be observed again here (based on FORCE, DISCRETE DISTANCE, PATH), the specific articulation of this movement through the quick succession of short tones (not a staccato anymore, but likely also based on the LINK schema), and, in a new development in this example, the additional rapid change of the direction of this movement (of either the whole group of animals, as in swarming, or of the animal’s individual wings, as in flutter and flight). This last underlying parameter could be based on what some authors in music cognition have called an OSCILLATION schema (Echard, 1999; Malawey, 2010). The emergent property in the blended space is again musical motion, yet articulated in a series of successive movements in opposite directions. For this reason, the participant can now interpret the music as “swarming”, “fluttering”, or creating a “flight path”. Again, the participants had to construct the experiential assumptions themselves in order to finalize the three richer images (and potentially many others). In Figure 3, these are allocated to the grounding box.

It seems, therefore, that constructs regularly employed by blending theorists, such as mental spaces, conceptual mapping, emergent structure, image-schema families, and grounding boxes can serve as good tools for postulating hypotheses on the reasons behind some typical descriptions of programmatic musical excerpts.

Conclusions

The goal of this article was to propose that the four-space model of the Conceptual Blending Theory could be successfully used in interpreting data from music conceptualization studies. The approach can be advantageous as it has sufficient theoretical tools to account for both questions of how new musical concepts emerge and how seemingly different concepts may derive from a similar set of constraints. The generic space contains the constraints, viewed as families of image schemas. Input space 1 incorporates the musical percept as a Gestalt, i.e. an instance of elementary conceptualization. Input space 2 hosts the referential description which the respondent draws from his or her personal, linguistic, or cultural circumstances. In the blend, we see how a novel conceptual property emerges from the system. Finally, semantically richer descriptions obtained in experiments require more contextual information, for which the notion of the grounding box may be useful in further work.

In essence, this proposal may be useful in that it (1) allows a lot of room for cross-cultural, cross-linguistic, and individual diversity in creating musical concepts, but also (2) provides the tools to capture the *motivation* behind the construction of such seemingly diversified responses.

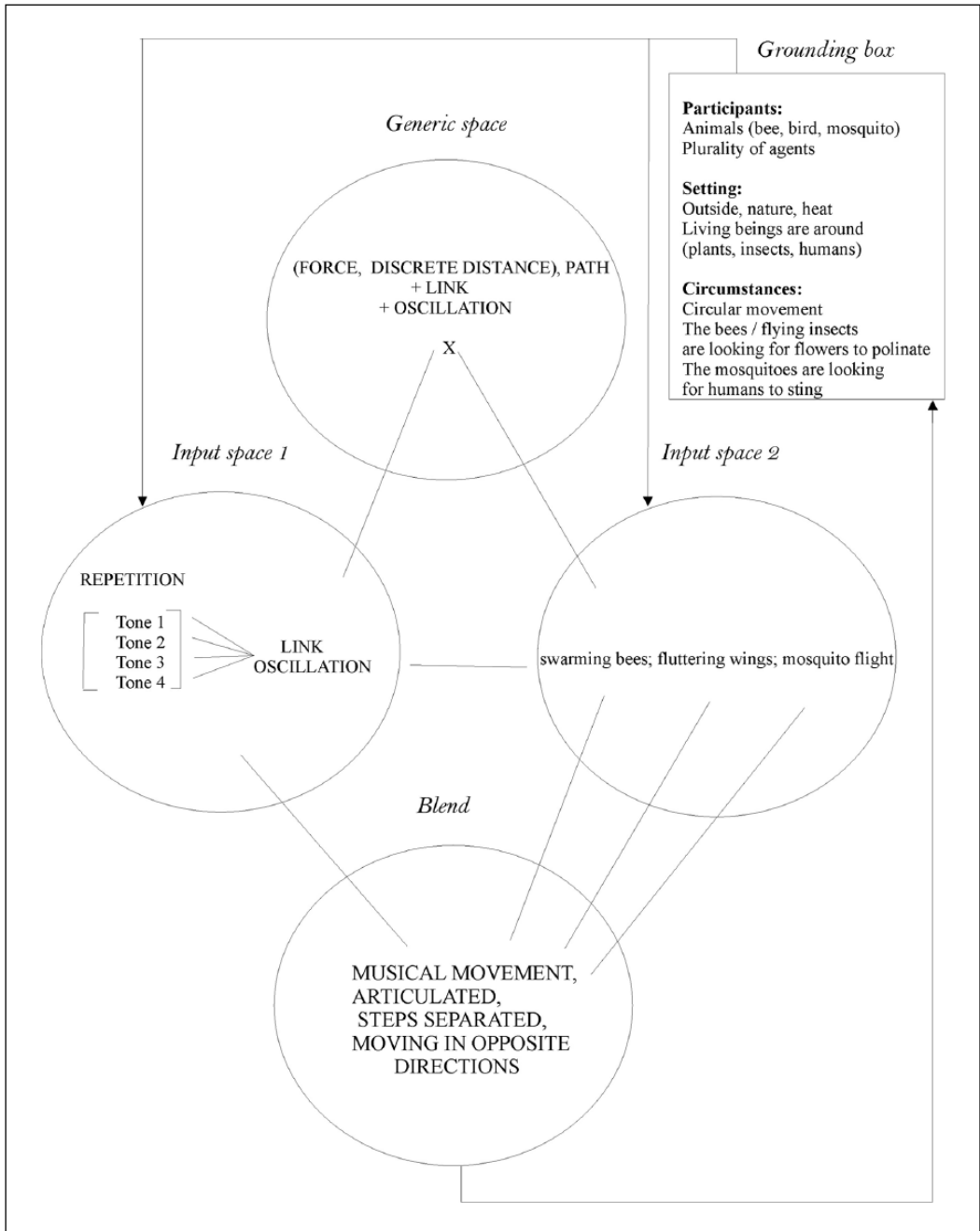


Figure 3. Conceptualization of the violin trills in Vivaldi's *Spring*.

As such, the approach might help deconstruct the sometimes too rigid epistemological distinctions in (musical) concept formation, such as the cleft between intramusical and extramusical meaning, or the disputes between universalism and relativism.

Further work should involve both theoretical and empirical efforts. Theoretically, the notion of the grounding box should be developed further, possibly no longer as a single construct, but as a multi-layered, hierarchical system incorporating biological, linguistic, and social contextual factors. Antović (2016) and Athanasopoulos and Antović (this issue) have now taken the first steps in this direction. Also, the approach proposed here would be more valuable if corroborated experimentally, and not used just as a means for post-hoc analysis. In this regard, Antović et al. (2017) have now obtained pilot results in which nonmusician participants seem to prefer various cross-culturally conditioned representations of pitch movement in a scale based on underlying postulated schemas, rather than the final form of the stimulus. More specifically, participants seem to strongly prefer a visual presentation of a musical scale in which both discrete movement and unidirectional path underlie the animation, irrespective of whether that animation reflects the concept in their mother tongue (a square going upward) or not (a square thinning in width or shrinking in all directions). So, the presence of strictly ordered image schemas in families seems more important to participants than their experiential, cultural knowledge of what the concept is normally “called” in their native language.

Another possible option for future work is to experiment on musical elements conceptualized in isolation, as opposed to those appearing in context. This might be a good way to study why, in the West at least, two adjacent tones are “close” to one another in a single-part melody, but indeed very “distant” in chords in a modulation, as of course keys are metaphorically “close” in terms of the number of pitches they have in common and not the proximity in frequency. Cross-cultural differences should also always be carefully taken into account. For but one example, the *FORCE* schema, discussed in the present work, may not have the same contextual implications in all communities. Reportedly, BaYaka Pygmies sing louder when putting their babies to sleep (Lewis, 2013), which may sound counterintuitive to Western listeners.

In effect, many interesting questions on conceptualizing music remain open. Hopefully, this article has suggested that an approach based on conceptual blending may be useful in such future endeavors.

Funding

This research was supported by the Serbian Ministry of Science (project number 179013).

References

- Antović, M. (2009). Musical metaphors in Serbian and Romani children: An empirical study. *Metaphor and Symbol*, 24(3), 184–202.
- Antović, M. (2010). From oceanic feeling to image schemata: Embodied mind and the construction of identity through binary conceptualization. In V. Lopičić & B. Mišić-Ilić (Eds.), *Identity issues: Literary and linguistic landscapes* (pp. 177–194). Newcastle, UK: Cambridge Scholars Publishing.
- Antović, M. (2014). Metafora o muzici ili metafora u muzici: jedan prilog za saradnju kognitivne lingvistike i kognitivne muzikologije [Metaphor about music or metaphor in music: A contribution to the cooperation of cognitive linguistics and cognitive musicology]. In M. Stanojević (Ed.), *Metafore koje istražujemo: suvremeni uvidi u konceptualnu metaforu* [Metaphors we study: Contemporary insights into conceptual metaphor] (pp. 233–254). Zagreb, Croatia: Srednja Europa. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2566258
- Antović, M. (2016). From expectation to concepts: Towards multilevel grounding in musical semantics. *Cognitive Semiotics*, 9(2), 105–138.
- Antović, M. (in press). Persuasion in musical multimedia: A Conceptual Blending Theory approach. In J. Pelcova & W. Lu (Eds.), *Persuasion in public discourse: Cognitive and functional perspectives*. Amsterdam, the Netherlands: John Benjamins.

- Antović, M., Bennett, A., & Turner, M. (2013). Running in circles or moving along lines: Conceptualization of musical elements in sighted and blind children. *Musicae Scientiae*, 17(2), 229–245.
- Antović, M., Mitić, J., & Benecasa, N. (2017). *Cross-modal interactions in the perception of pitch movement: Do language-independent primitives underlie the construction of musical concepts?* Manuscript submitted for publication.
- Antović, M., Stamenković, D., & Figar, V. (2016). Association of meaning in program music: On denotation, inherence, and onomatopoeia. *Music Perception*, 34(2), 243–248.
- Brandt, A., & Eagleman, D. (2017). *The runaway species: How human creativity remakes the world*. Edinburgh, UK: Canongate Books.
- Brandt, L., & Brandt, P. A. (2005). Making sense of a blend: A cognitive-semiotic approach to metaphor. *Annual Review of Cognitive Linguistics*, 3(1), 216–249.
- Brandt, P. A. (2008). Music and the abstract mind. *Journal of Music and Meaning*, 7. Retrieved from <http://www.musicandmeaning.net/issues/showArticle.php?artID=7.3>
- Brower, C. (2000). A cognitive theory of musical meaning. *Journal of Music Theory*, 44(2), 323–379.
- Cambouropoulos, E., Kaliakatsos-Papakostas, M., & Tsougras, C. (2015, August). *Structural blending of harmonic spaces: A computational approach*. Paper presented at the 9th Triennial Conference of the European Society for the Cognitive Science of Music (ESCOM), Manchester, UK. Retrieved from http://www.coinvent-project.eu/fileadmin/publications/3_structuralBlending_of_harmonicSpaces.pdf
- Chattah, J. (2006). *Semiotics, pragmatics, and metaphor in film music analysis* (Unpublished doctoral dissertation). Florida State University, FL.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, the Netherlands: Foris Publications.
- Cook, N. (2001). Theorizing musical meaning. *Music Theory Spectrum*, 23(2), 170–195.
- Coulson, S., & Oakley, T. (2005). Blending and coded meaning: Literal and figurative meaning in cognitive semantics. *Journal of Pragmatics*, 37(10), 1510–1536.
- Dolscheid, S., Shayan, S., Majid, A., & Casasanto, D. (2013). The thickness of musical pitch: Psychophysical evidence for linguistic relativity. *Psychological Science*, 24(5), 613–621.
- Echard, W. (1999). An analysis of Neil Young's "Powderfinger" based on Mark Johnson's image schemata. *Popular Music*, 18(1), 133–144.
- Eppe, M., Confalonieri, R., Maclean, E., Kaliakatsos, M., Cambouropoulos, E., Schorlemmer, M., & Kühnberger, K. U. (2015, July–August). *Computational invention of cadences and chord progressions by conceptual chord-blending*. Paper presented at the 24th International Joint Conference on Artificial Intelligence, Buenos Aires, Argentina. Retrieved from <https://www.ijaa.csic.es/en/node/54153>
- Evans, V. (2014). *The language myth: Why language is not an instinct*. Cambridge, UK: Cambridge University Press.
- Fauconnier, G., & Turner, M. (2002). *The way we think: Conceptual blending and the mind's hidden complexities*. New York: Basic Books.
- Grady, J., Oakley, T., & Coulson, S. (1999). Blending and metaphor. In R. Gibbs & G. Steen (Eds.), *Metaphor in cognitive linguistics* (pp. 101–124). Amsterdam, the Netherlands: John Benjamins.
- Hampe, B. (2005). Image schemas in cognitive linguistics: Introduction. In B. Hampe (Ed.), *From perception to meaning: Image schemas in cognitive linguistics* (pp. 1–14), Berlin, Germany: Walter DeGruyter.
- Hedblom, M., Kutz, O., & Neuhaus, F. (2015). Choosing the right path: Image schema theory as a foundation for concept invention. *Journal of Artificial General Intelligence*, 6(1), 21–54.
- Hedblom, M., Kutz, O., & Neuhaus, F. (2016). Image schemas in computational conceptual blending. *Cognitive Systems Research*, 39, 42–57.
- Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: The MIT Press.
- Johnson, M. (1987). *The body in the mind*. Chicago: The University of Chicago Press.
- Johnson, M., & Larson, S. (2003). "Something in the way she moves": Metaphors of musical motion. *Metaphor and Symbol*, 18(2), 63–84.
- Koelsch, S. (2012). *Brain and music*. Oxford, UK: John Wiley & Sons.
- Kramer, L. (2011). *Interpreting music*. Berkeley: University of California Press.
- Kühl, O. (2008). *Musical semantics*. Bern, Switzerland: Peter Lang.
- Lakoff, G. (1987). *Women, fire and dangerous things*. Chicago: The University of Chicago Press.

- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: The University of Chicago Press.
- Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York: Basic Books.
- Larson, S. (2012). *Musical forces: Motion, metaphor, and meaning in music*. Bloomington: Indiana University Press.
- Lewis, J. (2013). A cross-cultural perspective on the significance of music and dance to culture and society: Insight from BaYaka pygmies. In M. Arbib (Ed.), *Music, Language and the Brain* (pp. 46–65). Cambridge, MA: MIT Press.
- Malawey, V. (2010). Harmonic stasis and oscillation in Björk's *Medúlla*. *Music Theory Online*, 16(1). Retrieved from <http://www.mtosmt.org/issues/mto.10.16.1/mto.10.16.1.malawey.html>
- Mandler, J. M., & Pagán Cánovas, C. (2014). On defining image schemas. *Language and Cognition*, 6(4), 510–532.
- Pinker, S. (2013). *Language, cognition, and human nature: Selected articles*. New York: Oxford University Press.
- Saslaw, J. (1996). Forces, containers, and paths: The role of body-derived image schemas in the conceptualization of music. *Journal of Music Theory*, 40(2), 217–243.
- Sayrs, E. P. (2003). Narrative, metaphor, and conceptual blending in “The Hanging Tree”. *Music Theory Online*, 9(1). Retrieved from <http://www.mtosmt.org/issues/mto.03.9.1/mto.03.9.1.sayrs.html>
- Stefanou, D., & Cambouropoulos, E. (2015, August). *Enriching the blend: Creative extensions to conceptual blending in music*. Paper presented at the 9th Triennial Conference of the European Society for the Cognitive Science of Music (ESCOM), Manchester, UK. Retrieved from http://www.coinvent-project.eu/fileadmin/publications/2_enriching_the_Blend.pdf
- Swain, J. (1996). The range of musical semantics. *Journal of Aesthetics and Art Criticism*, 54(2), 135–152.
- Tasoudis, D., & Vouvaris, P. (2016, May). *B(l)ending time: Film music and meaning construction*. Paper presented at the 3rd Cognitive Science Workshop: Symbol Grounding in Cognitive Science, University of Niš, Serbia.
- Tsougras, K., & Stefanou, D. (2015, August). *Conceptual blending and meaning construction: A structural/hermeneutical analysis of the “Old Castle” from Mussorgsky's “Pictures at an Exhibition”*. Paper presented at the 9th Triennial Conference of the European Society for the Cognitive Science of Music (ESCOM), Manchester, UK. Retrieved from http://www.coinvent-project.eu/fileadmin/publications/1_conceptualBlending_and_meaningConstruction.pdf
- Watt, R., & Ash, R. (1998). A psychological investigation of meaning in music. *Musicae Scientiae*, 2(1), 33–53.
- Zacharakis, A., Kaliakatsos-Papakostas, M., & Cambouropoulos, E. (2015, October). *Conceptual blending in music cadences: A formal model and subjective evaluation*. Paper presented at the 16th International Conference on Music Information Retrieval, Malaga, Spain. Retrieved from http://www.coinvent-project.eu/fileadmin/publications/Zacharakis_et_al_ISMIR15.pdf
- Zbikowski, L. M. (2002). *Conceptualizing music: Cognitive structure, theory, and analysis*. New York: Oxford University Press.