Moving Figures and Grounds in music description

Wadley, Phillip; Tenbrink, Thora; Wallington, Alan

Cognitive Linguistics

18. Apr. 2024

Phillip Wadley*, Thora Tenbrink and Alan Wallington

Moving Figures and Grounds in music description

Abstract: This paper is a systematic investigation of motion expressions in programmatic music description. To address issues with defining the Source MOTION and the Target MUSIC, we utilize Gestalt models (Figure-Ground and Source-Path-Goal) while also critically examining the ontological complexity of the Target MUSIC. We also investigate music motion descriptions considering the role of the describer's perspective and communicative goals. As previous research has demonstrated, an attentional Goal-bias is common in physical motion description, yet this has been found also to lessen due to audience accommodation effects. We investigate whether this also occurs in music description. Using cognitive linguistic frameworks, we conducted an analysis of 21 English speakers' written descriptions of dynamic orchestral excerpts. All participants gave a description of one excerpt reporting their own personal experiences and the other excerpt reporting the events of the excerpt for a fictional future participant. We find that addressee accommodation shapes the choice of the ontological types of Figures used from being more subjective and creative in describing music for oneself versus being more objective in describing music for others. However, our investigation does not find sufficient evidence for a Goal-bias in music like there is in physical motion event descriptions.

Keywords: cognitive discourse analysis; Source-Path-Goal; music metaphor; Figure-Ground; Goal-bias; spatial language

1 Introduction

When people talk about 'motion' in music, what exactly do they mean? Johnson and Larson (2003) argue that our understanding of music is fundamentally metaphorical, allowing us to discuss music with language typically used for physical motion

*Corresponding author: Phillip Wadley, School of Arts, Culture, and Language, Bangor University, 39 College Road, Bangor LL57 2AP, UK, E-mail: phw18yvs@bangor.ac.uk
Thora Tenbrink and Alan Wallington, School of Arts, Culture, and Language, Bangor University, Bangor, UK, E-mail: t.tenbrink@bangor.ac.uk (T. Tenbrink), a.m.wallington@bangor.ac.uk (A. Wallington)

occurrences. Their position stems from linguistic evidence akin to the following examples:

(1)   "Nowhere in [Mahler's] music is there such an implacable, brutal <u>tread</u>."

(2)   "Then <u>comes</u> the famous 'Fate' motif."

(3)   "After a strange, otherworldly chorale, <u>this is followed by</u> its complete
        opposite – a lovely <u>surging</u> theme."

<div align="right">(Hewett 2010)</div>

These excerpts are from a critical description of the first minutes of Mahler's *6th Symphony.* The author uses motion words like *tread* (1) and *surging* (3) to depict a particular manner of motion; the idea that a motif *comes* (2) anchors to a deictic location; additionally, a *chorale followed by a theme* (3) shows that more than one thing can 'move' at a time along the same path. From these examples alone, we can see that movement is an important part of how people discuss music.

Music is temporal: it is a sound sequence that is not experienced in a single moment like an image would be. Rather, it manifests over a stretch of time, like a visual motion event, except with music, its events are not visually but aurally dynamic. Through Conceptual Metaphor analysis (Lakoff and Johnson 1980), scholars have been able to investigate patterns of such dynamic descriptions in the motion language of music discourse (Antović 2009; Julich 2018; Julich-Warpakowski 2019; Larson 1997; Pérez-Sobrino and Julich 2014), which has led to systematic insights into music cognition (Eitan and Granot 2006; Granot and Eitan 2011; Johnson and Larson 2003; Julich-Warpakowski 2022; Larson 1997; Zbikowski 1997, 2008, 2018). One central inquiry, however, in any examination of metaphor is how comprehensively and systematically the Source structure applies to the Target (Barnden and Wallington 2010; Gentner 1983; Lakoff 1993; Wallington 2010). Likewise, what is included in this Source-Target transfer from MOTION to MUSIC?

Within Conceptual Metaphor Theory, a cross-domain mapping describes a systematic relation between a Source and a Target domain, and such an approach could be used to explain broad aspects of metaphorical motion in music discourse. However, in general exactly what constitutes a domain is often vague, difficult to operationalize, and subject to discrepancies in inter-annotator agreement (Shutova et al. 2013), and we would suggest that the broad domains of both MUSIC and MOTION have this problem, being ambiguous as to what they should include. However, we shall argue that specific Gestalt models such as Source-Path-Goal and Figure-Ground (adapted into cognitive linguistics by scholars like Talmy (2000) possess none of the vagaries that pertain to the more general MUSIC and MOTION domains. They provide ample resources for establishing the conceptual structure of the Source domain for music metaphors, namely MOTION. Such integration of cognitive

linguistic frameworks provides a more comprehensive, nuanced, and replicable description of cognitive processes and mental representations in musical thought than an intuitive approach to defining the domain of MOTION. Similarly, MUSIC consists of a wide range of elements, from componental sounds, melodies, accompaniment, and sometimes narrative ideas, making it an equally complex Target. By considering MUSIC's ontological complexity, we can better understand what exactly is 'moving' in a music motion description.

One factor that has not been extensively examined in music motion research is communicative context, which we believe demonstrates Source-Target transfers. Particularly, these issues surround speaker perspective and communicative goals. One of our aims is to investigate how the same situational effects that occur in visual motion description, such as changes in attention, addressee adaptation, and relevance choices, arise in music motion description.

To the best of our knowledge, frameworks like Talmy's (2000) have not been directly utilized to explore movement configuration in music discourse, nor have situational context effects in music description been thoroughly investigated. So far, research in this area has revealed that how we think about music and how we talk about it are inextricably linked. In this empirical study, we investigate this complex relationship by methodically examining linguistic patterns using discourse analytic tools within cognitive linguistic frameworks. We aim to address two questions: firstly, what kinds of Figures and Grounds are used in music description and how are they integrated with the Source-Path-Goal schema and secondly, how does pragmatic context affect these descriptive choices?

First, we will outline the basic mechanics of spatial motion language from a Figure-Ground framework (Talmy (2000 [1975]). Radden and Dirven (2007) provide some details of the syntax-semantic interface which helps identify Sources, Paths, and Goals. In the next section, we will show how these linguistic configurations work in describing metaphorical movement in music description. This will lead us to the broader insights Conceptual Metaphor Theory (Lakoff and Johnson 1980) has offered in explaining why we use motion language to talk about music. We will also critique these insights' limitations. In the final part of our literature review, we will explore the importance of perspective and communicative aims in motion event descriptions. For our account of the present study, we will detail how we operationalized identifying linguistic patterns relative to the Source-Path-Goal schema alongside diverse kinds of motion Figures and situating Grounds that emerged from the data. After outlining the results, we will discuss what they mean for our research questions, their interpretive limits, and what our findings prompt for future study.

2 Background

2.1 Figure and Ground in motion description

Figures and Grounds are an elementary distinction in spatial language, adapted from Gestalt psychology (Koffka 1935) into cognitive linguistics first by Fillmore (1968) and later by Talmy (2000 [1975]). For the purposes of this paper, a Figure can be understood as an object or point that is "moving or conceptually movable" while a Ground functions as a reference point that situates that Figure (419). The Ground in contrast is comparatively more fixed in its location. Example 4 below illustrates these roles as expressed in language.

(4)　　The cat (F) jumped onto the bookshelf (G)

In English grammar, the Ground commonly comes in the form of a noun phrase that serves as a complement to a preposition (Radden and Dirven 2007). So, in this example, the Ground is the *bookshelf* as it informs us about where the cat is situated. Other information about the *cat's* location is implied or unimportant. One feature of these Grounds is they are attention-based. In (4), the person who formulated the sentence also had the choice of specifying where from and where through the cat jumped. This decision is determined by what Langacker (1987) refers to as *construal*, "the relationship between a speaker (or hearer) and a situation that he conceptualizes and portrays involving focal adjustments and imagery" (488). Speakers have choices in what they highlight and obscure when describing a motion event (c.f. Talmy 2000: 257–310). Choosing a bookshelf to situate a cat's movement was driven by the event's end point being the most important reference point for the person producing the utterance. The other options for a more complete motion event description could alternatively include:

(5)　　The cat jumped from the floor

(6)　　The cat jumped through the air

One way we can approach this choice is by analyzing the event as consisting of a Source-Path-Goal schema. The above examples (5) and (6) describe a Source and a Path, while example (4) describes a Goal. Prepositions themselves, such as *onto, through, by, from, to, towards, along, past*, and others can be thought of as directional adverbials, meaning they code an origination point, a route, or a destination point. These examples illustrate three types of prepositional phrases which express different roles for the Ground. The Source, Path, and Goal are typically entities (*floor, air, bookshelf*) preceded by directional or locative prepositions. Radden and

Dirven (2007)[1] provide a table of S-P-G-denoting prepositions that gives a general overview (310). We summarize them below:

> Source: (away) **from**, off (of), out (of), outside (of)…
> Path: **by,** past, via, along, about, around, through, throughout…
> Goal: **to**, at, for, towards, on(to), against, in(to)…

From, to, and *by* are described as the prototypical prepositional denotators for the Figure's motion route.

Such groupings together with the dissection of how these prepositions function, as Radden and Dirven (2007) present, provides a resource to begin to operationalize an analysis of the motion domain using prepositional phrases alone. Thus, music, examples such as the ones below can be approached much like (7–9):

(7)   The piece starts with blasts <u>from the horns</u> (Source)

(8)   The orchestra is playing <u>through some variations</u> (Path)

(9)   Finally, we get <u>to the coda</u> (Goal)

Here, this is a more abstract motion where the movement seems in some sense metaphorical (which we will examine in the next section), yet the analytical approach still works. In (7) *the horns* are a Source, *from* which a Figure *blasts* originates. *Some variations* serve as a Path-Ground in (8), as denoted by the preposition *through*. A contextually dependent Figure *we* in (9) moves *to* a *coda*, which is its Goal.

Note that we are not claiming that prepositional phrases constitute the sole means of expressing the Ground. We are merely claiming that they are a common way of doing so and that Radden and Dirven's (2007) table provides an existing resource that can be readily used in an analysis. For example, the notion of *lexical conflation* captures how certain movement verbs conflate Source, Path, Goal, and Manner (Talmy 2000 [1985]). *Arrive and enter,* for instance, conflate movement together with Goal, whilst *depart and exit* conflate movement with Source. As Talmy notes, languages differ in what they choose to conflate, with English more typically conflating Manner with motion than Source, Path, or Goal. Instead, the Source, Path, or Goal is expressed through a prepositional phrase or what Talmy terms a satellite. Partly for this reason, we feel justified in restricting our attention in this paper to prepositional phrases that express the Ground, leaving its expression through other means to future work. However, for music descriptions written in one of the Romance languages such as Italian, such an approach is likely to prove

---

1 Radden and Dirven use the terms TRAJECTOR and LANDMARK in their description rather than FIGURE and GROUND. For the purposes of this paper, we shall consider the terms as equivalent.

insufficient. There are other reasons for our decision to concentrate on preposi-
tions and that is if there is no preposition used to give the Ground, then it is
not always clear what constitutes the Ground. For example, in the sentence *George
returned*, Source and Goal Grounds are deictic and not identifiable out of context.
Moreover, verbs like *return* are ambiguous in identifying Grounds as Sources,
Paths, or Goals. The Source-Grounds could be either interpreted as George's
starting location (A) or his temporary secondary location (B), like in Example 10:

(10)   George returned [to his <u>office</u> (*A*) from the <u>breakroom</u> (*B*)].

Even when the Grounds are implied, quantifying something not mentioned does not
inform an analysis of attention in motion description, which is one of our aims.
Finally, whereas we can call on Radden and Dirven's list to give a quantitative
account of motion highlighting, to our knowledge, no standardized list of which
verbs conflate Source, Path, or Goal exists. There are ones that identify conflation of
Ground and Manner in verbs, such as that of Cifuentes Férez (2008), but specification
of more granular Ground (Source-Path-Goal) is more difficult for scholars to agree
upon. Alternatively, identifying explicit Grounds through preposition analysis is a
solid starting point for understanding how English speakers construe and configure
motion in music description.

Empirically, a phenomenon known as Goal-bias is a common finding in motion
description studies: usually, Goal-coding adverbials, including those coded as prep-
ositional motion constructions, are preferred over Source-coding ones. According to
scholars like Ikegami (1979: 148), a Goal is a more complete event when standing
alone than a Source. Taking our example, *the cat jumped onto the bookshelf*, it does
not pose the question of where it jumped from as strongly as *the cat jumped from the
floor* poses the question of where it jumped to.

There exists some debate on whether this is a cognitive bias (e.g., Verspoor et al.
1997) or a linguistic one (Stefanowitsch 2018). Stefanowitsch and Rohde (2004)
propose a *complete conceptualization hypothesis* whereby "we always attempt to
conceptualize motion events in their entirety (with equal attention to goal, [path],
and source)" (Stefanowitsch 2018: 145). However, cross-linguistically the phenom-
enon is widespread, from Greek (Johanson et al. 2019), Arabic (Regier and Zheng
2007), Japanese (Ihara and Fujita 2000), and English (Lakusta and Landau 2005;
Lakusta and Landau 2012; Papafragou 2010; Papafragou and Trueswell 2020), giving
credence to the cognitive-bias view. In English at least, the Goal-bias also manifests
by having a substantially longer list of different adverbials or prepositions in
corpus collocation studies like Stefanowitsch (2018: 164) whereas Source has a
relatively limited set, suggesting instead a linguistic bias.

Alternatively, there is also a strong case for Goal-bias having a pragmatic motivation, suggesting the Goal is generally the most relevant spatial reference point for describing a motion event at the time of utterance. Relevance Theory (Sperber and Wilson 1986) argues that human minds are geared toward relevance in discourse, this may offer an account of why Goals are most highlighted: identifying an endpoint to a motion event has most inference potential for identifying the Path and Source. If one says they are *going to the kitchen*, then the Source should be implied from the context (the immediate vicinity) and the Path can be inferred by which route they would probably take between those two points. This pragmatic view veers closer to *complete conceptualization hypothesis* proposed by Stefanowitsch and Rohde (2004) but diverges by only claiming that the granular details of information about the motion event is accessible not automatically, but only when deliberated upon.

With music, we want to know if there exists such a Goal-bias in motion description. However, this is not certain as music has a highly complex ontology and this could have a bearing on what is designated to the roles of Figure and Ground.

2.2 Metaphorical motion in music

The motion language used in music discourse has been a subject of interest for researchers ever since the advent of Conceptual Metaphor Theory (Lakoff and Johnson 1980). Johnson and Larson (2003) wrote a landmark paper that established motion as a key component of music cognition by drawing parallels to motion metaphors of TIME. Like Figures and Grounds in TIME, metaphorical motion in music depicts different possible conceptualizations of what entities move and what entities are stationary. With this, we can begin asking the question: which ontological entities most often play the respective roles of Figure or Ground in music description?

Johnson and Larson propose three kinds of metaphorical motion in music: these are MOVING MUSIC, MUSICAL LANDSCAPE, and MOVING FORCE. In their own words, they reason that "(a) we see objects move, (b) we move our bodies, and (c) we feel our bodies moved by forces" (2003: 68–69). MOVING MUSIC and MUSICAL LANDSCAPE are converse configurations of one another. The MOVING MUSIC examples they use include (69):

(11)   Here comes the recapitulation

(12)   the strings slow down now

(13)   the music goes faster here

According to Johnson and Larson, the musical 'objects' in these cases move relative to the Ego; they are "a musical event […] moving towards and then past us."

Alternatively, MUSICAL LANDSCAPE includes (71):

(14)   We're coming to the coda

(15)   the soloist is waiting to come in seven measures from here

(16)   the melody rises up ahead

In the first two examples they give, the music is a space in which a mobile Figure – performer, audience, or another musical entity – can 'move' around. Notably, (16) does not intuitively fit into MUSICAL LANDSCAPE, but perhaps more into MOVING MUSIC, though could be thought of as an instance of fictive motion (c.f. Julich 2018).

The authors base MOVING MUSIC and MUSICAL LANDSCAPE on the MOVING TIME and MOVING EGO temporal schematizations proposed by Clark (1973). Like MOVING MUSIC, MOVING TIME involves an Ego reference point for the movement of TIME (for which MUSIC seems to be a substitute). MUSICAL LANDSCAPE and MOVING EGO conversely share a structure where the Ego moves relative to a stationary TIME-Ground. To use a simple analogy by interchanging TIME with MUSIC, the parallel is clear:

(17)   The [deadline/coda] is approaching. (MOVING TIME/MOVING MUSIC)

(18)   We're approaching the [deadline/coda]. (MOVING EGO/MUSICAL LANDSCAPE)

Lastly, MOVING FORCE is when both the Ego and the music move, specifically when the latter acts as a cause of motion for the former: *the music moved me emotionally* or *the piece took the audience on an adventure.* As Dąbek (2020: 196) observes, MOVING FORCE has no clear equivalent to Clark's (1973) TIME schematizations.

Johnson and Larson's (2003) proposed metaphors have a certain highlighting quality for Figures and Grounds: MOVING MUSIC emphasizes the Figure, whatever a 'musical entity' includes, whereas MOVING LANDSCAPE emphasizes the Ground, evoking the image of a landscape to move across. By studying their examples, we find that MOVING LANDSCAPE depicts a relatively fixed set of locations (*the coda, seven measures from here, ahead*) in contrast to musical entities which are conceptually mobile (*we, the soloist, the melody*). Both Figures and Grounds are relatively easy to identify in moving LANDSCAPE, where while Ground is highlighted, Figure is not necessarily obscured. This could be because Figures are harder to omit from a motion description than a Ground would be, at least in English. For example, in *the orchestra rushed through the movement*, one can ask which can best stand on its own: *the orchestra* (F) *rushed* or *through the movement* (G)?

While Johnson and Larson's (2003) framework has provided an insightful start to dissect the structure of music motion, there are a number of problems. Firstly, it is debatable whether all their example sentences do illustrate the mapping they

claim. Sentences 12 and 13 are argued to show music/strings moving relative to an Ego at the deictic center. However, they could alternatively be analyzed as describing a *slowing down* or *going faster* relative to their own speed at an earlier period. The issue here is not which analysis is better but how variable and replicable such analyses should be. Secondly, their approach is not holistic and, as some scholars have pointed out, overgeneralizes what is a much more complex set of mappings. One start to this is Pérez-Sobrino and Julich's (2014: 306) Music Event Structure Metaphor (MESM), based on Lakoff's (1993) Event Structure Metaphor for dynamic events. While the MESM does not exclusively include motion metaphors in music description, it does elaborate more on the comparatively broad mappings proposed by Johnson and Larson (2003) through corpus analysis. Table 1 shows the first four parts.

The first three categories of the MESM follow Johnson and Larson's (2003) three aforementioned types but with more specification. (I) MUSICAL ENTITIES ARE LOCATIONS ON A LANDSCAPE evokes the landscape image but does not specify the Figure identity, musical or non-musical. (II) MUSICAL PROGRESSION IS MOTION ON A LANDSCAPE/MUSIC IS A JOURNEY is an elaboration on the first part, from a stationary spatial description to a motion description. Notably, one divergence of Pérez-Sobrino and Julich's (2014) iteration is that the moving Figure is not limited to only the Ego here, as its time metaphor basis MOVING EGO suggests. However, Johnson and Larson (2003) are a little unclear on whether the EGO always plays the role of the moving Figure. (III) MUSICAL CHANGE IS FORCED MOTION (DUE TO MUSICAL CAUSES) broadly covers causal relations between musical elements, a specification of Johnson and Larson's MOVING FORCE. The other categories of the MESM are more culturally specific, such as with the inclusion of (IV) UNEXPECTED HARMONIES ARE IMPEDIMENTS TO MOTION ALONG THE PATH but

Table 1: Event Structure Metaphor conversion to Music Event Structure Metaphor.

Event Structure Metaphor (ESM) Lakoff (1993)	Music Event Structure Metaphor (MESM) Pérez-Sobrino and Julich (2014)
"STATES ARE LOCATIONS"	→ (I) "MUSICAL ENTITIES ARE LOCATIONS ON A LANDSCAPE"
"PURPOSEFUL ACTION IS SELF-PROPELLED MOTION"	→ (II) "MUSICAL PROGRESSION IS MOTION ON A LANDSCAPE/MUSIC IS A JOURNEY"
"CAUSES ARE FORCES, CAUSATION IS FORCED MOTION"	→ (III) "MUSICAL CHANGE IS FORCED MOTION (DUE TO MUSICAL CAUSES)"
"DIFFICULTIES ARE IMPEDIMENTS TO MOTION"	→ (IV) "UNEXPECTED HARMONIES ARE IMPEDIMENTS TO MOTION ALONG A PATH"

still evokes a Figure-Ground relation in terms of MOTION (Pérez-Sobrino and Julich 2014: 306).[2]

The MESM further hints at different possible ontological identities of Figures and Grounds. Grounds seem the most straightforward: as a LANDSCAPE broadly corresponds to a Ground, it can be specified further with IMPEDIMENTS, DESTINATIONS, and LOCATIONS, which all seem to be instances of Goals and Paths (though, no Sources can be found in the MESM). Unlike Johnson and Larson's (2003) metaphors, the linguistic annotations of Pérez-Sobrino and Julich (2014) show examples of MUSICAL PROGRESSION IS MOTION ON THE MUSICAL LANDSCAPE in a wide array of musical 'spaces', including not only temporal progression, but also harmonic progression, and pitch change. This indicates a wider ontological diversity of Grounds in music description than Johnson and Larson (2003) propose.

Johnson and Larson (2003) argue that TIME IS MOTION motivates the music motion metaphors they propose. However, Julich (2018: 145), in a later study, demonstrates that this is an oversimplification. For example, harmony can be conceptualized spatially as a CLOSE-REMOTE relation where a point in a melody can be located relative to a *home key* or *tonal center.* In a description of a harmonic relation, this can then be expressed in language in examples like:

(19)   we can trace the general harmonic *motion from* tonic *to* dominant

(20)   He [Lohengrin] will *return to* a flat-key area

Compare with this non-musical example from (Wallington 2015: 32):

(21)   The names are in alphabetical order, so 'Barnes' should *come before* 'Brown'.

(19) and (20) can be conceptualised without TIME involved in the mapping. (21) is an example of a non-temporal use of the SEQUENCE IS RELATIVE POSITION IN A PATH metaphor. MOTION, however, is still metaphorically present in both examples (19)–(20) and (21) where Figures and Grounds are as well clearly identifiable. In (19), it is describing a relation between *we* and *motion* (as co-Figures) and the *tonic* (Source) to the *dominant* (Goal); (20) depicts *Lohengrin* (Figure) as moving to a *flat-key area* (Goal); (21) describes the name *Barnes* (Figure) moving relative to *Brown* (Path) while *Brown* also acts as an additional moving Figure. As Julich (2018) argues, complex motion events in music have much basis specifically in SEQUENCE rather than more broadly in TIME.

She points to Evans's (2003) argument that TIME has no single experiential structure and instead covers a wide array of different experiences, such as SEQUENCE,

---

2 Other phenomena of course could be metaphorically thought of as 'impediments' other than unexpected harmonies, such as irregular rhythms or sudden changes in volume. The inclusion of this is owed however to the prevalence it has in the authors' corpus.

CHANGE, and DURATION. The same can be said about music. What this suggests, as Julich (2018: 137) elaborates, is "Instead, motion metaphors present a complex group of metaphors that employ different kinds of spatial organization and motion to conceptualize various aspects of the Target domain of music." Hence, the supposed transfer from TIME IS MOTION to MOVING MUSIC and MOVING LANDSCAPE overgeneralizes MUSIC as a Target domain. Instead, Julich lists six spatial sub-mappings she found in her corpus study that includes the Targets TEMPORAL PROGRESSION, POSITIONS IN THE MUSICAL PIECE, SEQUENTIAL ORDER, TEMPO, DURATION, and SIMULTANEITY (146).

These Targets are all possible contributors to different ontological identities for Figure-Ground. Her findings we think prompt addressing a central ambiguity of Johnson and Larson's (2003) paper and the MESM: what is a 'musical entity'? It can be many possible objects or ideas: it is not clear whether narrative, non-musical sound, and temporal terms are musical entities, for instance. Hence, a broad definition for a 'musical entity' encompassing each of these runs into essentially the same issue as the Target domain TIME. Greater specification is therefore needed for more meaningful insights, and Julich (2018) identifying more specific Targets in music is a useful start to this.

Figure-Ground ontology is especially interesting to consider as evidence exists that different components of the Target domain MUSIC can have their own unique metaphorical cognition. Timbre, pitch, rhythm, and tempo, i.e., the aural building blocks of music have been shown to have metaphorical mappings in studies by Eitan and Granot (2006) and Granot and Eitan (2011). In their experiments, they found that some of these mappings were asymmetrical, where "musical change in one direction evokes a significantly stronger spatial analogy than its opposite" (Eitan and Granot 2006: 221). One example of this is the analogy between dynamics and directionality: "while diminuendi are strongly associated with descending spatial motion, crescendi are not significantly associated with ascents" (235).[3] On a more basic sound level as well, Grassi et al. (2013), Lemaitre et al. (2017), and Rusconi et al. (2006) have conducted studies on non-musical sound and find complex systems of spatial representations from different combinations of sounds. Several diverse kinds of entities move in music on just the aural level, and there is no reason to assume that these entities have the same motion behaviors and emphases.

Source domain diversity as well supplies complexity to music description. In particular, music historically has also been often paired with narrative, such as in opera, ballet, television, and films, which provides rich mental images at the imagination's disposal to associate, making this a major cultural factor to consider. For example, a popular musical technique known as 'Mickey-Mousing' could play a

---

3 Diminuendi are decreases in volume with crescendi are increases.

significant role in how dynamic programmatic music is processed. The term is used by the film industry to describe beat-by-beat parallels of musical gestures and visual events, often in musical scores for cartoons such as *Looney Tunes* or *Tom and Jerry*. For running up a staircase on screen, the score might utilize an upwards chromatic scale, or a character being hit with something might be accompanied by a cymbal crash (Bradley 2002). While this alignment has been practiced in multi-modal musical performance long before cartoons, mainly from pantomime and melodrama (Shapiro 1984), cartoon music as some musicologists argue (Audissino 2020; Brown 2011; Goldmark 2011) has greatly influenced popular Western conceptions of what kinds of events we interpret the music to be representing. Musical gestures that parallel visual events could be reified to an extent where they stand on their own as representing those visual events. Eitan and Granot (2006) report similar icon-driven reification in their experiments on musical sound perception. As a formal area of research, this has received very little attention from metaphor analysis.

What these considerations suggest is a mix of different ontological categories of Figures and Grounds in music description and music cognition. Most importantly this interface – including rich blends (Zbikowski 2008, 2018) – makes the ontology of music discourse unique from other types of motion description. As demonstrated, MUSIC is a broad Target domain, and much more multifaceted than how Johnson and Larson (2003) model it. The diversity of Targets for music prompts the question of referentiality: what exactly is being referred to by MUSIC? It could include the instruments (which are usually metonymies for the performer or sound produced (Johnson and Larson 2003: 76), the sequence of events, the theoretical components (*scales*, *chords*, etc.) the emotions evoked, or the narrative troupes signaled.

From componential sounds to emergent musical ideas, along with the multifaceted blends of temporal language which has extensive scholarly coverage of its own (Evans 2003; Fauconnier and Turner 2008; Wallington 2010), and – on top of this – the emergence of narrative meaning, all of these contribute to a highly complex domain for MUSIC to which MOTION can apply in a plethora of ways.

2.3 Situational and communicative considerations in music motion description

So far, we have looked at how the Target of MUSIC is referentially multi-faceted: many different components of varying ontologies are involved. We also hold that gestalt structures like Figure-Ground and S-P-G could be used in structuring the Source for motion metaphors in general. Music discourse has yet to undergo such an examination. However, motion description always is unavoidably pragmatic, giving yet greater complexity to the Source MOTION. As the central argument of Conceptual

Metaphor Theory (Lakoff and Johnson 1980) states that we not only talk about but also *experience* concepts in terms of other concepts, situational aspects of such experiences like perspective and relevance could also play an integral role. From this, we have two questions: *where does the speaker then situate themselves relative to the motion event? And how does the communicative purpose shape the description?*

Concerning the first question, in one setting, a person can be a participant as either the Figure moving (*you walk through the park*) or as a Ground marker that is relatively stationary to the moving object (*the dog walks past you*). In another setting, the person could be a bystander to a motion event where they are observing it from the outside (*the man walked past the dog*). These are egocentric and allocentric spatial reference frames, respectively, and based on these, the choice of Figure and Ground identity may be affected.

Johnson and Larson's (2003) framework, like Clark's (1973) temporal schematizations, exhibits largely an egocentric model of how we experience music. The way in which they characterize each metaphor, MOVING MUSIC, MOVING LANDSCAPE, and MOVING FORCE, all involve the assumption of the Ego being part of the motion event as either a Figure or a Ground. However, Julich-Warpakowski (2022, 2018) and Peréz-Sobrino and Julich (2014) both show a striking absence of Ego in their corpus data of music criticism. This is puzzling since music description for participants in rehearsal, for instance, can be egocentric, like in the following examples:

(22)   *you* (F) are supposed to come in at Figure 24 (Path)

(23)   then *we* (F) go into the key change (Goal)

(24)   once the melody (F) comes to *you* (Goal), be sure to match the style

So, if instances such as these do exist in music discourse, the absence of Ego in some contexts could just be situational. A possible factor influencing this is agency: in the context of a conductor saying any of examples (22)–(24) to musicians in an orchestra, they are essentially foregrounding the ensemble's role for the outcome of the performance by situating them in a metaphorical space where the musicians interact with the elements of the music. Unlike a music critic or audience member who is a passive listener, the performers are active participants in its production, so it makes sense that the EGO would be more involved.

Allocentric motion description has interesting empirical trends as well which could potentially carry over to music motion description, specifically in Ground attention. Returning to the phenomenon of Goal-bias, this tendency can lessen when the describer assumes the audience is unfamiliar with the Source identity as demonstrated by Do et al. (2022), Lakusta and Landau (2012) and Papafragou and Trueswell (2020). This prompts some very interesting questions. For example, could Path mentions also be affected by the audience when treated as its own reference

point type, whether this Goal-bias transfers also to metaphorical motion in music, and if the accommodational shift in motion description also occurs in music description?

The consideration of addressee knowledge brings us to the second pragmatic inquiry of ours: how communicative purpose affects music motion description. Describing music for oneself will be different from how one would describe it for someone else, as music is at least cognitively a very personal and idiosyncratic experience. On the other hand, describing the same experience for someone else will adhere to principles of relevance (Sperber and Wilson 1986), deeming some details important while others are discarded.

One research area which has shown a difference in controlling for communicative purpose in spatial language description is wayfinding. For example, Hölscher et al. (2011) in one of their urban route description studies placed participants in groups of For-Self and For-Other addressee conditions when asked to describe a prospective route from one location to another. The For-Self participant group was asked, "Which route would you take if you now wanted to walk from here to the [next destination]?" (2011: 237). The For-Other group was asked, "Imagine that somebody else who is not familiar with this area has to walk from here to the [next destination]. Please consider now which route this person should take." The researchers found greater detail in 'for-other' route descriptions, more salient landmarks used as reference points, and fewer turns. From studies such as these, the purpose of the description whether it is for a prescribed route for an addressee or if it is a description of an individual's personal route demonstrates a pragmatic effect on the content of participants' spatial language: accommodation to expected mutual knowledge. Tenbrink (2012) emphasizes that such studies highlight the importance of relevance in spatial language usage.

Applying this to music, Reybrouck (2015: 6) argues that the listening strategies involved in actively attempting to make meaning of music resemble that of wayfinding descriptions, involving "working out a *cognitive map* or *route-description* that is the outcome of the structuring of previous and actual experiences". This includes deictic "mental pointing", stepwise presentation of information, and mental simulation of travel. By mental pointing, the author is referring to primarily deictic terms: *here/there, come/go, now/then,* etc. The use of these can be quite telling of both the extent of mutual knowledge assumptions as well as how they situate themselves and the addressee in the description. Stepwise presentation of information has a sequential quality to it which, like giving instructional steps to an addressee, music description is formulated in a way that guides the addressee in their prospective experience. Mental simulation of travel is particularly interesting in terms of music considering the ambiguous role of fictive motion. In route description, fictive motion is the use of motion describers for stationary locations. For example, *the trail runs*

over the mountain is not a literal description of the trail's behavior, but as some argue (Talmy 2000 [1996]) is a mental simulation of movement of the speaker or addressee along that path. With music, the integral role of fictive motion has been addressed by Julich (2018: 178) in instances like *Measures 3–4 traverse the same space as measures 1–2*. In physical route description, examples like *the road goes from the clocktower to the park* likewise suggest some motion already metaphorized in a physical context, adding an extra layer of complexity in the interface between the two domains of MUSIC and MOTION.

It should be emphasized though that route description in wayfinding has pragmatic differences to music description. As far as communicative purpose, a listener does not have any control over what occurs in the music they are listening to, unlike someone giving directions in wayfinding. The only exception might be a person creating the music or playing it. This should have a bearing on the spatial language choices used as the communicative purpose is descriptive rather than prescriptive. In addition, wayfinding studies like Hölscher et al. (2011) are predisposed to egocentric descriptions of routes as participants were asked to plan a route for themselves or someone else, and with that agency, the use of first and second-person prepositions are more common. For music, many other possible Figures than Ego are at the speaker's disposal with its high ontological complexity. Finally, no wayfinding studies we are aware of utilize the S-P-G schema to describe the prospective, situated, or retrospective motion involved. Studies on allocentric motion (ex. Lakusta and Landau 2012; Papafragou and Trueswell 2020; Stefanowitsch and Rohde 2004) description do, however, provide at least a basis of generally what to expect in a for-self condition where there is minimal pragmatic accommodation in motion description.

3 Methodology

3.1 Cognitive discourse analysis

The present study uses cognitive discourse analysis (CODA) methods. As the literature we described above integrates three wide-ranging research areas – motion description, music metaphor, and pragmatics – we believe an exploratory approach to our research question is most appropriate. Our aims are descriptive since we have no precise predictions.

CODA is a methodology formalized by Tenbrink (2020) that integrates cognitive linguistic frameworks with discourse analysis. The aim of CODA is to investigate complex thought processes and mental representations through systematic analysis of the linguistic features of discourse data, collected in controlled situations where

the speakers express their thoughts. CODA is particularly good for such studies as its analytic scope is broad, considering both cognitive and pragmatic aspects. It also has fruitfully been used in spatial cognition studies (Hölscher et al. 2011; Tenbrink and Salwiczek 2016; Tenbrink and Seifert 2011). CODA aims to not only analyze what is expressed in language, but also *how* those thoughts are expressed, as the precise formulation can reveal underlying patterns of concepts – as in the case of motion expressions in music, and other metaphors used in everyday language.

Our study aims to address two main factors affecting motion expressions in music. The first one is ontological: what kinds of Figures and Grounds in the broad Target domain of MUSIC do people use in music description? Investigating this, we will examine what categories which emerge from the data we gather with consideration of what was explored in Section 2.2.

The second question is a pragmatic one: how does the purpose of and the addressee of a description affect these choices? Here, we investigate how relevance and mutual knowledge play a role in music motion description as we described in Section 2.3 as well as attention to Source-Path-Goal entities as described in Section 2.1.

3.2 Procedure

Our study was presented in the form of an online survey that included a required consent form at the beginning, links to two musical stimuli, an open response for each excerpt, and a background questionnaire. The survey was shared on online forums and sub-forums on reddit.com such as r/classicalmusic and r/samplesize. As we were interested in discourse by proficient English speakers only, participants were asked to rate for themselves their language proficiency on a scale from 1 (poor) to 5 (native level). We only used data of people rating themselves 4 or higher.

For the musical stimuli, we selected two short Stravinsky excerpts (Craft 2005a; Craft 2005b): *Petrushka (Tableau III: The Blackamoor (2:26))* (excerpt 1)[4] and *The Nightingale (Scene 3: Prelude (2:43))* (excerpt 2).[5] Both are programmatic, written for moving characters, the first for a ballet and the second for an opera. Each is purely instrumental, involving no human voices, and the titles of the pieces were not disclosed to the participants. Participants were only provided links to mp3 audio files of the excerpts with no visual supplement of neither the live performances nor the album covers. The reason we chose these excerpts is for their dense eventfulness which resembles a 'Micky-Mousing' style we described earlier. This specific type of

---

4 Track 27 from: Craft (2005a).

5 Track 33 from: Craft (2005b).

music provides ample opportunity for participants to imagine narratives, attach meaning to intensifying and abating musical stimuli, and it does not require specialized knowledge of music theory concepts. Mainly, we aimed to make the stimuli accessible to all participants.

Following a similar procedure to Hölscher et al. (2011), we prompted participants to describe the music with the main distinction being whether the description is meant to reflect their own experience or to describe it for someone else. *In as much detail as possible, please describe the events in the music from beginning to end. What happens? What changes?* followed by a free-text space for their response. In the For-Self (FS) condition, participants were asked to *describe the music as you personally experienced it* while in the For-Other (FO) condition they were asked to do so *for a future participant about to listen to it.* Each participant listened to both musical excerpts and was asked to describe one of them for the FS audience and the other for the FO other audience. There were four versions resulting from alternating the two addressee conditions as well as switching the order of the excerpts, namely:

A: (Excerpt 1 + For Self, Excerpt 2 + For Other)
B: (Excerpt 2 + For Other, Excerpt 1 + For Self)
C: (Excerpt 1 + For Other, Excerpt 2 + For Self)
D: (Excerpt 2 + For Self, Excerpt 1 + For Other)

We sent out four versions of the study comprising of different excerpt orders and audience combinations to assure that any effects we observe are not due to one excerpt providing more descriptive opportunities than the other. From this measure, we can collapse across all For-Self descriptions and al For-Other descriptions in our analysis.

After the open responses, participants took a background questionnaire asking for age and gender, as well as their musical education and classical music exposure. This included a) how often they listen to classical music, b) whether they play a classical instrument and how often, and c) what level of theory education they have had.

Twenty-one participants took part in this study, distributed as follows: A: ($n = 5$), B: ($n = 6$), C: ($n = 5$), and D: ($n = 5$). Seven of the participants were male, 13 were female, and 1 preferred not to say. Nineteen of the 21 participants were between the ages of 18–55, with the remaining two in the 56+ range. Nine reported playing an orchestral instrument regularly, 12 listened to classical music every day or every other day, and 15 had either some basic or advanced-level music theory education.

Of our participants, 9 can thus be confidently categorized as musicians, whereas others had musical knowledge to varying extents. It stands to reason that participants volunteered to contribute to this study based (at least in part) on their interest in music. As our study does not set out to ask whether musicians think about music differently than non-musicians, the differences in music expertise should not affect

our results concerning addressee effects, because all participants responded to both addressee registers. As studies such as Tenbrink et al. (2019) and Cialone et al. (2018) show, specialists tend to give more detailed descriptions, so musicians making up a large population of our study should provide us with a richer dataset than if comprised of only non-musicians.

3.3 Source-Path-Goal analysis

Our analysis in determining the preferences for using Sources, Paths, or Goals in situating descriptions focused on grammatical evidence rather than lexical evidence. We used Radden and Dirven's (2007: 310) table of orientating prepositions for this approach, which we reference in Section 2.1.

All instances of prepositions in the S-P-G list above we termed as 'prepositional motion constructions': prepositions that indicate an orientated motion occurring with one of the prepositions listed above followed by a Ground entity. By counting the occurrence of these prepositions in our data, we can compare S-P-G highlighting in participants' descriptions and be able to identify which Ground entity entities are S-P-G.

These prepositions in their descriptions are usually clear by attaching to a specific Ground, as it often comes immediately after the preposition, making them easy to count. Two cases serve as exceptions to this construction: One is where prepositional phrases do not have a clear Ground, which some grammarians (Huddleston and Geoffrey 2006; Pullum 2009) refer to as 'intransitive prepositions'. For example, *they got away* or *he went further* both only consist of a motion trajectory. For our analysis, examining what subjects explicitly mention tells us more about salience and attention than implied information does. We still included these cases under the subcategory Intransitive (I) along with S-P-G to gauge the amount of implicit information the participants assume their audience knows.

The other exception to the standard constructions we looked for is where a Goal is mentioned with no preposition preceding it, like in *we made it home*. Pullum (2009) makes an argument that *home* is functionally an intransitive preposition in cases like these. However, *home* has more specific semantic content than typical intransitive prepositions, so we counted it as both a prepositional motion phrase and a Goal.

Notably, many of the empirical studies we have mentioned analyze S-P-G differently from Radden and Dirven's (2007) prepositional approach. For example, Do et al. (2022) only focus on Source and Goal in their analysis while similarly Lakusta and Landau (2012) only code Path as attached to a Source or Goal, namely as Source-Paths (ex. *out of the house)* and Goal-Paths (ex. *into the house*). Stefanowitsch

(2018) on the other hand uses the term Trajectory to refer to a similar class of motion adverbials to Radden and Dirven's (2007) list for Path. We decided to use Radden and Dirven's list as it is more standardized than other lists and because we want to test whether Path mentions are any more prominent in music motion description than physical motion description studies.

3.4 Figure and Ground analysis

Figure and Ground identification was the second part of the analysis. Semantic roles, as we pointed out, can inform us of the mental representations people have in their minds when listening to and describing music. We primarily want to know what the Figures and Grounds are: are they instruments, melodies, affects, the audience, or something else? Julich (2018) has laid the foundations for identifying musical entities in greater granularity than MUSIC and EGO. The Figure and Ground types in our analysis foremost emerge from the data. Importantly, our approach is informed by two forces: the categories must be broad, but not overlap in any of its members (Krippendorff 2004; Tenbrink 2020).

For counting an entity as a Figure, it must be accompanied by 1) a prepositional motion construction, 2) an intransitive preposition, or 3) a lexical motion term. A 'lexical motion term' is a construction surrounding a verb, adjective, or adverb (or a verb, adjective, or adverb by itself) which describes a motion behavior, including orientational, cyclical, or self-contained motion. For example, in *the clarinet rushed*, the verb *rushed* is the lexical motion term. In Talmy's (2000 [1985]) terms, the verb conflates a specific Manner of motion, but not necessarily require a Path satellite (ex. *through* the passage). We counted lexical motion phrases so that Figures can be identified easily since they often accompany one another as prepositions do with Grounds. Lexical motion phrases sometimes occur in conjunction with prepositional motion phrases. In these cases, we counted them as separate motion indicators. For example, *the clarinet rushed through a passage* would count *rushed* as a lexical motion phrase, *through a passage* as a prepositional motion phrase (*passage* as the Path-Ground entity), and *the clarinet* as the Figure. Verb valency plays an important part in this, as some motion verbs do require a complement phrase to complete it, meaning that some adverbials, prepositions, and objects may be more syntactically distant from the head verb than others. The identification of Figures and Grounds thus requires looking at descriptive clauses and sentences as a whole.

Lastly, our analysis of Figures and Grounds approaches motion as it is explicitly described. For example, we do acknowledge that Huumo (2017) makes a compelling

case that EGO, when acting as a Ground for verbs like *return,* it is often implied. Such cases can similarly be made about lexical conflation and other aspects as well. However, as we stated earlier, we are limiting our analysis to be as non-inferential as possible since it is more difficult to quantify implied information. This means we are counting individual occurrences of Figures and Grounds as they appear in text which should inform us about what the describers deem important to highlight.

4　Results

The results listed below are quantitative findings. In the first part of the analysis, we will give a general overview of the data we collected before illustrating what we found with Source-Path-Goal constructions. We will then elucidate the identities of these Grounds alongside the Figure identities (i.e., instruments, musicological terms, narrative concepts, etc.) For each, we will introduce a typology that emerged from the data.

4.1　General quantitative findings

Table 2 provides an overview of the data we collected.

In total, we examined a dataset of 4,052 words with 125 prepositional motion constructions (PMCs), 199 Figures, and 173 lexical motion terms (LMTs).

In the For-Other condition, the average word count (WC) is higher, indicating a greater degree of elaboration in descriptions of this register. This measure is the total of all participants' WCs divided by the number of participants within each condition. A two-tailed paired sample *t*-test indicated a significant difference between the amount of elaboration in which individual participants engaged when they were describing music for themselves or for others ($t(20) = -2.36$, $p = 0.0286$).

Table 2: Overview.

Participants	$N = 21$				
	Total WC	**Average WC**	**PMCs**	**Figures**	**LMTs**
For-Self	1,819	86.62	46	83	71
For-Other	2,233	106.33	79	116	102
Total	4,052		125	199	173

4.1.1 Prepositional Grounds and S-P-G

For examining attention in prepositional motion constructions, Table 3 illustrates some trends.

Out of 125 prepositional motion constructions we identified, 39 of them were Intransitive. From this table, we can see that around a third of prepositional motion phrases in both addressee conditions do not anchor to a specific reference point. Examples of these include: "the rest of the orchestra joins *in*," "the two notes come *back* again," "as the piece goes *on*," and "tension section, like flying or waving *around*."

Another trend is the general lack of Path Grounds in our dataset (FS = 1, FO = 5). Source and Goal on the other hand are more prevalent and additionally appear more contrasted in their emphasis. The immediate suggestion this data makes is a general tendency for the omission of Source in the FS condition, appearing as a Goal-bias, however, this is not statistically significant. To test whether the contrast of Source-Goal mentions has a relationship to the situational difference between FS and FO, we employ a Chi-Squared test of independence. We omitted Path since the number of observations does not meet the minimum of 5 per condition for this test. Since we are interested in the distribution of reference points, we did not include Intransitive as they are Grounds with no entity referent attached. We found 80 observations of Source or Goal in our dataset, but only 6 of Path. Using the Chi-squared goodness of fit test comparing the proportion of only the Source and Goal entities out of these observations between For-Self and For-Other (as Chi Squared tests require a 5-count minimum which Path comes short in both groups), the result is not significant (χ^2(1, N = 21) = 2.45, p = 0.118). Even with a Fisher test including Path, the result showed similar insignificance (p = 0.237). Nonetheless, the numerical difference with the discrepancy on the end of FS may repay further investigation with a larger sample.

Table 3: Attentional distribution – Source-Path-Goal.

Motion grounds	FS (rf)	FO (rf)	FS (*n*)	FO (*n*)	Total
Source	19.6 %	31.6 %	9	25	34
Path	2.2 %	6.3 %	1	5	6
Goal	43.5 %	32.9 %	20	26	46
Intransitive	34.8 %	29.1 %	16	23	39
Total			46	79	125

4.1.2 Figure and Ground categories in music descriptions

The number and criteria of Figure and Ground categories emerged from the data we analyzed. We began by intuitively dividing different entities that are used as reference points that situate Figures in descriptions and the Grounds themselves, such as distinguishing instruments from basic sounds from theory terms, etc. In the process, we created word lists and descriptive criteria which we show below in full:

SEQUENTIAL – References to order and sequence
   Moment, beginning, end, before, after, earlier, the next second, destination, return, peak
SECTIONAL – chunked sections of the music referred to in general, non-musical language
   part, snippet, section, passage, , episode, get-go, line, excerpt, material, bit, bars
INSTRUMENTAL – any specific instrument or instrument group named
   Instruments, horn, strings, string section, the cello, violas, brass, winds, cymbal, woodwinds, orchestra, violin, trumpet, violin section, , cor anglais, percussion, flute, piccolo, bassoon, the rest of the orchestra, low instruments, solo, English horn, bells
MUSICOLOGICAL – music theory, musicological, and performance terms
   melody, note, forte, tune, waltz, harmony, tremolo, grace note, arpeggios, an airy, fanfares, phrase, series of low notes, runs, minor key, motif, idea, trills, pizzicato, suspension, minor sounds, loud chord, set of notes, the theme, staccato, whole notes, a pronounced rhythm, flourish, block chords, interval, lick,
TRAIT – general descriptors of qualities which the music goes to or from
   louder, much, quiet, black, complexity, completion, volume, stature
NARRATIVE – constructed characters or story-based entities
   someone, something, conversation, a clock, alert, discovery, realization, they, calls, festival, hunter, prey, people, assertion, the other members, reason, crawling, interruptions, goblin, loud interjections, thunder
AFFECTIVE – emotions and psychological states
   feeling, anxiety, intensity, tension, impression, dread, sensitivity, confusion, agitation
GENERIC – broad reference to the music as a whole or terms that are too broad for other categories
   the music, it, together, group, originator, piece, string entries
SOUND – basic units of non-musical sound, pitch, and timbre
   crash, bang, echoes, low (pitch), high (pitch), pitch, sound, tones, plucks, above, descent, layers
EGO – use of the first or second person or reference to a listener
   You, I, we, the listener, my attention

Some of these terms are not immediately intuitive as Figures or Grounds by themselves: for example, *quiet, before, black, above,* and *much* are not usually considered entities. However, each of these examples are functionally used as Grounds in their grammatical contexts in the data we analyzed: "drop to *quiet,*" "Oboes quote the bassoon theme from *before,*" "it ends with a cheery fade to *black,*" "each with a grace not from *above,*" "[the music] doesn't build to *much.*" In each example, they are points of reference for describing a dynamic motion event even though they are not nominal entities.

From this typology, we categorized the Figures and Grounds and calculated the proportion in which they appear in FS and FO datasets. Tables 4 and 5 illustrate the proportions for Figures by category.

The first striking detail we find is the overall lack of *ego* Figures in both conditions (FS (n) = 4, FO (n) = 2) as shown in Table 4. *Instrumental* Figures were the most dominant (FS (rf) = 26.5 %, FO (rf) = 27.6 %) and not contrasted in situational usage, with *generic*, *musicological*, and *narrative* following as the most common categories of moving entities. One interesting point here is how rare *sequential* Figures are, in both conditions only being used once.

To determine if there is an addressee effect, we again utilized a chi-squared test of independence on categories that had more than 5 observations in both conditions. Certain categories contribute to the Chi-Squared value more than others, as the total of each category's contribution scores add up to the total Chi-Squared value of 17.57. We calculated the Chi Squared contribution by squaring the difference between the observed and expected values for each and dividing that by the expected value. The expected values themselves are the products of each count of category observations

Table 4: Figures – categorical analysis.

Category	FS (rf)	FO (rf)	FS (n)	FO (n)	Total
instrumental	26.5 %	27.6 %	22	32	54
narrative	22.9 %	6.0 %	19	7	26
musicological	13.3 %	31.9 %	11	37	48
generic	12.0 %	12.1 %	10	14	24
affectual	7.2 %	6.9 %	6	8	14
sectional	6.0 %	6.0 %	5	7	12
sound	6.0 %	6.9 %	5	8	13
ego	4.8 %	1.7 %	4	2	6
sequential	1.2 %	0.9 %	1	1	2
			83	**116**	**199**

Table 5: Figures – Chi-squared test of independence analysis.

Obs. (*n*)	DF	Chi-Sq	*p* value
199	5	17.63	0.0035

	Figure	Observed	Contribution to χ^2 test
For-Self			
	narrative	19	6.39
	musicological	11	3.91
	sectional	5	0.05
	affectual	6	0.01
	instrumental	22	0.00
	generic	10	0.00
	Total	**73**	
For-Other			
	narrative	7	4.48
	musicological	37	2.74
	sectional	6	0.03
	affectual	8	0.01
	instrumental	32	0.00
	generic	14	0.00
	Total	**104**	

and the audience condition total divided by the grand total of all observations for both conditions. This was done for each number of category observations. Hence, a category's contribution to the Chi Squared test in both conditions indicates how much it is affected by the addressee. We find that there indeed is a relationship (χ^2(5, *n* = 21) = 17.63, *p* = 0.0035) where participants used fewer *narrative* Figures and more *musicological* Figures when describing the music for others than describing it for themselves.

Table 6 shows the same analysis as Table 4 but for Ground identities, i.e., entities included in a transitive PMC (*n* = 85). Once again, the *instrumental* category was dominant and is the only category that has more than the 5-observation minimum needed for a Chi-Squared test of independence. For this reason, we performed another Fisher test, which came out non-significant (*p* = 0.0886). However, the numerical difference (FS (*n*) = 6, FO (*n*) = 23) does provide some evidence of addressee accommodation, also when considering a similar trend occurs with Figure identities, suggesting further investigation with a larger dataset is warranted before drawing any definite conclusion.

Table 6: Transitive Grounds – categorical analysis.

Ground	FS (rf)	FO (rf)	FS (*n*)	FO (*n*)	Total
sequential	20 %	5 %	6	3	9
instrumental	20 %	42 %	6	23	29
sectional	13 %	11 %	4	6	10
musicological	13 %	13 %	4	7	11
trait	10 %	11 %	3	6	9
generic	10 %	2 %	3	1	4
affective	7 %	2 %	2	1	3
narrative	3 %	2 %	1	1	2
sound	3 %	17 %	1	7	8
Total			**30**	**55**	**85**

5 Discussion

Our study set out to examine Figure-Ground and situational effects in motion descriptions of programmatic music. Our results mainly account for the outcomes of the contextual conditions we tested, i.e.: in the written mode, one-way in communication, and aiming to describe a complex event. This will naturally be quite different from a spoken, two-way exchange discussing a simple event (Bell 1984). The key difference between our study's context and the context of music criticism (e.g., Pérez-Sobrino and Julich 2014) is primarily in the aim of the description. In music criticism, the aim is to evaluate the quality of a piece of music or performance of it. In our study, the participants described the events of the music and/or how it made them feel. Based upon differences observed in which Figures were employed in motion descriptions, we have found evidence that the addressee (even when that addressee is fictional) affects the degree of subjectivity and specialized knowledge with which participants describe music. By subjectivity, we mean the use of personal imagined entities as referents which could be ambiguous to others. In addition, like in Hölscher et al. (2011), participants elaborated on the music's events in greater detail and less ambiguously when describing it for an external addressee than when describing their personal experience. Alongside this key finding, the absence of evidence for a Goal-bias in the results overall suggests that music motion description is not subject to the same attentional biases as physical motion description. Beyond this overview of the results, we can see some more specific trends.

The first is that there are very few *ego* Figures, as some of the literature (Julich 2018; Julich-Warpakowski 2022; Pérez-Sobrino and Julich 2014) suggested would be the case, and contrary to the metaphors proposed by Johnson and Larson (2003).

Deictic mental pointing, like what Reybrouck (2015) describes when we actively try to make meaning of music, does not have a strong presence as far as the EGO is concerned. Returning to the dataset we gathered, we found the use of the word *there* often, but it was only used in its existential sense rather than pointing to a specific entity. Appearances of verbs *come* (FS (n) = 3, FO (n) = 7), *go* (FS (n) = 0, FO (n) = 5, and *return* (FS (n) = 4, FO (n) = 0) did indicate some deictic use though with no clear difference between conditions. These findings appear to be evidence of a preference for an allocentric spatial frame being used by participants overall. They rarely place themselves (or proxy selves, like *you, the listener,* etc.) as partakers in the motion events they describe. This demonstrates that wayfinding is not a strongly comparable model for music description in the context of being a non-participant in the production of the music.

Secondly, we find a disparity between the number of Figures (n = 199) and Grounds (n = 80). However, this is not substantially owed to multiple Figures anchoring to a single entity Ground, as this only occurs twice in the dataset. Instead, the occurrence of multiple Figures in a single motion event is most often attached to only a motion verb with no Ground, in examples like:

(25)　　the third part *brings back* the alternating horns and winds

(26)　　Finally, the violins, then violas as well, *pick up* the melody with quiet and unsettling long tones.

(27)　　A sombre, moody string line is *intercepted* by a soft percussion and energetic flute/piccolo

There are 25 such constructions in the data we gathered, encompassing 29 % of all Figures (n = 58) we identified. Taking the Johnson and Larson (2003) framework, this indicates the greater prevalence of the MOVING MUSIC metaphor over MOVING LANDSCAPE in our dataset.

Following Julich (2018), MUSIC as a Figure can be taxonomized meaningfully across 9 ontological referent types. We found that some of these are affected by contextual conditions, giving evidence specifically to the uniqueness of *musicological* and *narrative* Figure motion behaviors. As demonstrated by their differences between contexts and their contributions to their subsequent tests of independence, these two attest to a degree of pragmatic appropriateness depending on the communicative goal. This tendency was not exhibited strongly for Ground types. Yet, a visible numerical difference between conditions in the use of *instrumental* Grounds suggests a possible contextual effect where *instruments* act as Grounds more often in describing the music for others.

In an interesting contrast, while uses of *musicological* and *narrative* Figures experienced contextual effects, uses of *instrumental* Figures are relatively

unchanging. One reason this could be is because *musicological* and *narrative* Figures are more emergent agents than instruments, meaning that they have properties that consist of more than the sum of their parts. For example, a *melody* is made up of notes, dynamics, and rhythms in a certain structure but can be recognized as being melancholy, agitated, or joyful. This makes such Figures subjective to a degree that some listeners might not conceptualize it the same way as others. Instruments, on the other hand, while having some indirectness to their sign structure (i.e., instruments are metonyms for their sound (Johnson and Larson 2003)), are more componential in the music's events, making them reliably unambiguous as Figures.

A similar explanation could connect to the higher frequency of *instrumental* Grounds in For-Other descriptions. The most common preposition for these was 'from' where an *instrumental* Ground serves as an origin point for a Figure:

(28)   a short pizzicato flourish <u>from the strings</u>

(29)   an aggressive repeated high-low interval <u>from the horn section</u>

(30)   to a soft round sound <u>from a brass instrument</u>

This constructional configuration accounts for 68 % of FO Sources and 33 % of FS Sources that we identified. Knowing this, instruments seem to work quite well as Ground referents when describing music for another person since they are recognizable and clear in their reference, much like how it is as a Figure. It is especially intriguing to see that instruments can work as Figures or Grounds, though, as the former, *instrumental* Figures do not commonly anchor to a Ground entity, and instead their motion is most often indicated by a motion verb, like *enter*, *come*, or *leave*. This is often accompanied by an Intransitive null Ground:

(31)   the horns *come* <u>back</u>

(32)   different instruments *coming* <u>together</u>

(33)   *bringing* <u>in</u> the piano

S-P-G highlighting also exhibits a couple of interesting aspects. One is the meagre appearance of Path highlighting. Just like the studies of Lakusta and Landau (2012), Papafragou and Trueswell (2020) and Stefanowitsch (2018), Path does not warrant much attention for quantitative analysis in music as well, especially with our dataset. The question of whether this is a cognitive bias or a linguistic one is still up for debate. In English, despite its range of different Path adverbials overshadowing Source adverbials, Path adverbials are not as prominently included in physical motion description as Source or Goal, and this seems to be the same case for metaphorical motion in music. The second general finding here is that Intransitive adverbials are just as prolific as Source and Goal as it makes up a third of prepositional motion

constructions. Music description is highly directional, yet much of its motion is not anchored to specific reference points.

In addition, though the numerical results were suggestive of a Goal-bias in the FS condition, the disparity between Source and Goal is not statistically significant. Resultantly, we cannot confirm an addressee accommodation effect like what is found in a physical motion event description (Papafragou and Trueswell 2020). This could be explained in a couple of ways. One is that the experimental design did not sufficiently encourage participants to describe the music for another person, as the instructions read "In as much detail as possible, please describe the events in the music from beginning to end for a future participant about to listen to it. What happens? What changes?" Along with this, directly prompting participants to describe the music in terms of movement and motion may have benefited the richness of the data. A larger dataset would in addition give greater confidence to this result.

The second explanation for this could be interpreted as a finding applicable to music metaphor overall: music motion description can be considered reified to a degree that it is more rigid in its grammatical structure than literal motion description. Diegnan (2005) similarly demonstrates that metaphors are often constrained in their grammatical form. Hence, this might yield a negligible difference in prepositional indicators of motion for our own dataset. For example, in our data, the word *attack* is only used as a noun describing the start of a note, often followed by a Source referent, in all cases being the instrument:

(34)   An abrupt *attack* <u>from the brass</u>

However, *attacks* do not seem to go *to* or *through* any Ground referents. This could be interpreted as evidence against the Stefanowitsch and Rohde (2004) *complete conceptualization hypothesis,* as *attacks* do not seem to have a conventionally verbalizable motion schema which includes a Path or Goal. It is difficult additionally to imagine what Paths or Goals an *attack from the brass* could have. Either these motion schemata are only partially adapted in music metaphor or grammatical convention in music metaphor has no real preference for Source, Path, or Goal.[6] The main result we can say was affected by contextual conditions is Figure identity, suggesting, unlike grammatical formulation, lexical choices have relative flexibility for different contextual aims and audiences. The intrapersonal for-self condition yielded more narrative descriptions that were highly idiosyncratic. Take, for example, an excerpt

---

6 Interestingly, nothing like *[the instrument] attacks [x]* appears. Though, for musicians, this may be different as one might *attack a note* by playing it suddenly, however *attacking* another instrument or non-musicological object is not a regular use of the term, nor is the passive form (**the note was attacked*).

of the For-Self description (35) by one of our participants in contrast to their For-Other description (36):

(35)   At the beginning, there was a horn melody that evoked the feeling of a chase, like on horseback. Then, a slow, vibrating, soft sound created a feeling almost like being watched, or stalked. I felt like I was trying to hide from someone or something.

(36)   In the first part, loud brassy horns alternate with quieter woodwind instruments giving the listener a sensation of anxiety as the volume of the piece changes rapidly and suddenly. The whiplash from the quiet volume to the louder volume was intense.

This participant demonstrates a switch in the type of entities involved in their descriptions, from *chases, being on horseback, hiding*, and *being watched* and then instead to *instruments*, *melodies*, and *quiet volume to the loud volume*. The general use of more conventional referents and events, including non-motion ones, shows that participants are mindful to a degree of the subjectivity of their experience listening to the music. Overall, participants lean toward these creative descriptions more often when describing music for themselves while adjusting their focuses to more mutually accessible referents when describing music for others.

Two themes broader than this though have emerged from our findings. The first is the confirmation that an analysis of music motion description can be operationalized systematically while using the same approaches linguists use for physical motion description, something which we set out to demonstrate in this paper. Clearly delineating the components and structure of a motion event according to motion models like Talmy (2000) and Radden and Dirven (2007), brings highly granular insights to the extent of the Source-Target transfer between MOTION and MUSIC that would not be reached with an introspective cross-domain analysis. This brings us directly to the second theme that applies to metaphor mappings: they are often much more complex when placed under such systematic investigation. Music metaphors, as we find here, are mapped according to their specific communicative situation, making the broader mapping proposals of Johnson and Larson (2003) overgeneralizations. Instead, we find a much richer and more complex set of sub-mappings that are utilized to varying degrees depending on communicative purpose.

For future directions, examining wider datasets under the same analytical approach can bring crucial insights. Namely, the presence of a Goal-bias and accommodation effects would be important findings either way in better understanding how systematic Source-Target transfers are in metaphor. For questions of accommodation and situational effects, the possibilities are wide ranging. These include whether musicians and non-musicians construe music motion differently or

how players and listeners situate themselves relative to the music. While music is a highly individual experience, we do have the tools to identify underlying patterns of how the experience is conveyed through language.

6 Conclusions

Our study set out to explore Figure-Ground identity and contextual effects in music description with consideration of previous studies in physical motion description and music metaphor. Motion in music description has been previously identified as metaphorical and our contribution sought to examine how the analogues of music and motion map, as well as what the limitations of those mappings are. Our contribution most critically provides a demonstration of the utility of specifying both MUSIC and MOTION with the aim of a robust and systematic operationalization. For MOTION, we use gestalt models like the Figure-Ground and S-P-G frameworks for analyzing metaphorical motion just as it would be used for physical motion analysis. For MUSIC, we find that it is multi-faceted, and its diverse components emerge from the data we analyzed. In our study, we find insufficient evidence of Goal-bias when describing music for oneself, unlike what is found in physical motion description literature. However, we do think further research into this with a wider dataset is needed. In addition, we find that *musicological*, *narrative*, and *instrumental* Figures and Grounds have distinct motion behaviors which are affected by audience accommodation, mutual knowledge assumptions, and communicative goals.

Data availability statement: The collected dataset, including its metadata, annotation and analysis, is available on the OSF depository permalink https://osf.io/2gmcy/.

References

Antović, Mihailo. 2009. Musical metaphors in Serbian and Romani children: An empirical study. *Metaphor and Symbol* 24(3). 184–202.

Audissino, Emilio. 2020. The function of Mickey-Mousing: A re-assessment. *Sound and image: Aesthetics and practices*, 145–160. London: Focal Press.

Barnden, John & Alan Wallington. 2010. In Armin Burkhardt & Brigitte Nerlich (eds.), *Tropical truth (s): The epistemology of metaphor and other tropes*, 85–121. Germany: Verlag Max Neimeyer.

Bell, Allan. 1984. Language style as audience design. *Language in Society* 13(2). 145–204.

Bradley, Scott. 2002. Music in cartoons. In Daniel Goldmark & Yuval Taylor (eds.), *The cartoon music book*, 115–120. Chicago: Review Press.

Brown, Gregory Scott. 2011. *Carnaval, carny, and cartoons: Carl Stalling's cartoon music in historical context*. Salt Lake City: The University of Utah.

Cialone, Claudia, Thora Tenbrink & Hugo J. Spiers. 2018. Sculptors, architects, and painters conceive of depicted spaces differently. *Cognitive Science* 42(2). 524–553.

Cifuentes Férez, Paula. 2008. *Motion in English and Spanish: A perspective from cognitive linguistics, typology and psycholinguistics*. Universidad de Murcia PhD thesis.

Clark, Herbert H. 1973. Space, time, semantics, and the child. *Cognitive development and acquisition of language*, 27–63. Cambridge: Academic Press.

Craft, Robert & The Philharmonia Orchestra. 2005a. *Stravinsky, I.: Petrushka: Tableau III: The blackamoor. The robert craft collection: Stravinsky, I.: The Firebird/Petrushka. 3*. London: Naxos.

Craft, Robert & The Philharmonia Orchestra. 2005b. *Stravinsky, I.:The nightingale: Scene 3 (The Emperor's Bedchamber): Prelude. The robert craft collection: Stravinksy, I.: The rite of Spring/The Nightingale. 3*. London: Naxos.

Dąbek, Karolina. 2020. The metaphor of movement and its materialisation in twentieth-century spatial music. *Kwartalnik Młodych Muzykologów UJ. Koło Naukowe Studentów Muzykologii UJ* 4(47). 177–195.

Diegnan, Alice. 2005. *Metaphor and corpus linguistics*. Amsterdam: John Benjamins.

Do, Monica Lynn, Anna Papafragou & John Trueswell. 2022. Encoding motion events during language production: Effects of audience design and conceptual salience. *Cognitive Science* 46(1). e13077.

Eitan, Zohar & Roni Y. Granot. 2006. How music moves: Musical parameters and listeners images of motion. *Music Perception* 23(3). 221–248.

Evans, Vyvyan. 2003. *The structure of time: Language, meaning and temporal cognition*. Amsterdam: John Benjamins.

Fauconnier, Gilles & Mark B. Turner. 2008. In Raymond Gibbs Jr. (ed.), *Rethinking metaphor*. New York: Cambridge University Press.

Fillmore, Charles J. 1968. Lexical entries for verbs. *Foundations of Language* 4(4). 373–393.

Gentner, Dedre. 1983. Structure-mapping: A theoretical framework for analogy. *Cognitive Science* 7(2). 155–170.

Goldmark, Daniel. 2011. Sounds funny/funny sounds: Theorizing cartoon music. In *Funny pictures: Animation and comedy in studio-era hollywood*, 257–271. Berkeley and Los Angeles: University of California Press.

Granot, Roni Y. & Zohar Eitan. 2011. Musical tension and the interaction of dynamic auditory parameters. *Music Perception* 28(3). 219–246.

Grassi, Massimo, Massimiliano Pastore & Guillaume Lemaitre. 2013. Looking at the world with your ears: How do we get the size of an object from its sound? *Acta Psychologica* 143(1). 96–104.

Hewett, Ian. 2010. A masterpiece driven by sinister visions, marital love and cowbells. *The Telegraph*. Available at: https://www.telegraph.co.uk/culture/music/classicalmusic/7904110/Mahler-Year-Symphony-No-6.html.

Hölscher, Christoph, Thora Tenbrink & Jan M. Wiener. 2011. Would you follow your own route description? Cognitive strategies in urban route planning. *Cognition* 121(2). 228–247.

Huddleston, Rodney & Pullum Geoffrey. 2006. Coordination and subordination. In *The handbook of English linguistics*, 198–219. Hoboken: Wiley Online Library.

Huumo, Tuomas. 2017. The grammar of temporal motion: A Cognitive Grammar account of motion metaphors of time. *Cognitive Linguistics* 28(1). 1–43.

Ihara, Hiroko & Ikuyo Fujita. 2000. A cognitive approach to errors in case marking in Japanese agrammatism: The priority of the goal-ni over the source-kara. *Constructions in cognitive linguistics*, 123. Amsterdam: John Benjamins.

Ikegami, Yoshihiko. 1979. "GOAL'OVER'SOURCE": A case of dissymetry in linguistic orientation. *Angol Filológiai Tanulmányok/Hungarian Studies in English* 12. 139–157.

Johanson, Megan, Stathis Selimis & Anna Papafragou. 2019. The source-goal asymmetry in spatial language: Language-general vs. language-specific aspects. *Language, Cognition and Neuroscience* 34(7). 826–840.

Johnson, Mark L. & Steve Larson. 2003. "Something in the way she moves"-metaphors of musical motion. *Metaphor and Symbol* 18(2). 63–84.

Julich, Nina. 2018. Metaphors for musical motion – beyond time is motion. In Annalisa Baicchi, Rémi Digonnet & Jodi L. Sandford (eds.), *Sensory perceptions in language, embodiment and epistemology* (Studies in Applied Philosophy, Epistemology and Rational Ethics), vol. 42, 133–153. Cham: Springer International Publishing.

Julich-Warpakowski, Nina. 2019. Chapter 9. Why do we understand music as moving? The metaphorical basis of musical motion revisited. In Laura J. Speed, Carolyn O'Meara, Lila San Roque & Asifa Majid (eds.), *Converging evidence in language and communication research*, vol. 19, 165–184. Amsterdam: John Benjamins Publishing Company.

Julich-Warpakowski, Nina. 2022. *Motion Metaphors in Music Criticism: An empirical investigation of their conceptual motivation and their metaphoricity*. Amsterdam: John Benjamins.

Koffka, Kurt. 1935. *Principles of Gestalt psychology*. London: Routledge.

Krippendorff, Klaus. 2004. *Content analysis: An introduction to its methodology*, 2nd edn. Thousand Oaks, Calif: Sage.

Lakoff, George. 1993. The contemporary theory of metaphor. In Andrew Ortony (ed.), *Metaphor and thought*, 202–251. New York: Cambridge University Press.

Lakoff, George & Mark Johnson. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.

Lakusta, Laura & Barbara Landau. 2005. Starting at the end: The importance of goals in spatial language. *Cognition* 96(1). 1–33.

Lakusta, Laura & Barbara Landau. 2012. Language and memory for motion events: Origins of the asymmetry between source and goal paths. *Cognitive Science* 36(3). 517–544.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar: Theoretical prerequisites*, vol. 1. Stanford: Stanford University Press.

Larson, Steve. 1997. Musical forces and melodic patterns. *Theory and Practice* 22. 55–71.

Lemaitre, Guillaume, Hugo Scurto, Jules Françoise, Frédéric Bevilacqua, Olivier Houix & Patrick Susini. 2017. Rising tones and rustling noises: Metaphors in gestural depictions of sounds. *PLoS One* 12(7). e0181786.

Papafragou, Anna. 2010. Source-goal asymmetries in motion representation: Implications for language production and comprehension. *Cognitive Science* 34(6). 1064–1092.

Papafragou, Anna & John Trueswell. 2020. Cognitive and pragmatic factors in language production: Evidence from source-goal motion events. *Cognition* 205. 1–15, 104447.

Pérez-Sobrino, Paula & Nina Julich. 2014. Let's talk music: A corpus-based account of musical motion. *Metaphor and Symbol* 29(4). 298–315.

Pullum, Geoffrey K. 2009. Lexical categorization in English dictionaries and traditional grammars. *Zeitschrift für Anglistik und Amerikanistik* 57(3). 255–273.

Radden, Günter & René Dirven. 2007. *Cognitive English grammar*. Amsterdam: John Benjamins.

Regier, Terry & Mingyu Zheng. 2007. Attention to endpoints: A cross-linguistic constraint on spatial meaning. *Cognitive Science* 31(4). 705–719.

Reybrouck, Mark. 2015. Real-time listening and the act of mental pointing: Deictic and indexical claims. *Music, Mind, and Language* 2. 1–17.

Rusconi, Elena, Bonnie Kwan, Bruno L. Giordano, Carlo Umilta & Brian Butterworth. 2006. Spatial representation of pitch height: The SMARC effect. *Cognition* 99(2). 113–129.

Shapiro, Anne Dhu. 1984. Action music in American pantomime and melodrama, 1730–1913. *American Music* 2(4). 49–72.

Shutova, Ekaterina, Barry J. Devereux & Anna Korhonen. 2013. Conceptual metaphor theory meets the data: A corpus-based human annotation study. *Language Resources and Evaluation* 47. 1261–1284.

Sperber, Dan & Deirdre Wilson. 1986. *Relevance: Communication and cognition*. Oxford: Blackwells.

Stefanowitsch, Anatol. 2018. The goal bias revisited: A collostructional approach. *Yearbook of the German Cognitive Linguistics Association* 6(1). 143–166.

Stefanowitsch, Anatol & Ada Rohde. 2004. The goal bias in the encoding of motion events. In Günter Radden & Klaus-Ulwe Panther (eds.), *Studies in linguistic motivation*, 249–268. Berlin: Mouton de Gruyter.

Talmy, Leonard. 2000. *Toward a cognitive semantics: Concept structuring systems*, vol. 1. Cambridge: MIT Press.

Tenbrink, Thora. 2012. Relevance in spatial navigation and communication. In Cyrill Stachniss, Kerstin Schill & David Uttal (eds.), *Spatial cognition VIII* (Lecture Notes in Computer Science), vol. 7463, 358–377. Berlin, Heidelberg: Springer.

Tenbrink, Thora. 2020. *Cognitive discourse analysis: An introduction*. Cambridge: Cambridge University Press.

Tenbrink, Thora, Ruth C. Dalton & Anwen Jago Williams. 2019. *The language of architectural diagrams. Application/pdf*, 14. Wadern/Saarbruecken, Germany: Schloss Dagstuhl – Leibniz-Zentrum fuer Informatik GmbH.

Tenbrink, Thora & Lucie H. Salwiczek. 2016. Orientation and metacognition in virtual space. *Journal of Experimental Psychology: Human Perception and Performance* 42(5). 683–705.

Tenbrink, Thora & Inessa Seifert. 2011. Conceptual layers and strategies in tour planning. *Cognitive Processing* 12(1). 109–125.

Verspoor, Marjolijn H., René Dirven & Günter Radden. 1997. Putting concepts together: Syntax. *Introduction to language and linguistics: A cognitive approach*, 89–116. Duisburg: Gerhard-Mercator-Universität/Gesamthochschule.

Wallington, Alan. 2015. Uncertain futures: What light can metaphor shed upon the conceptualisation of time? In Emmanuel Labeau & Qiaochao Zhang (eds.), *Taming the TAME systems*, 25–38. Leiden: Brill.

Wallington, Alan M. 2010. 11. Systematicity in metaphor and the use of invariant mappings. In Graham Low, Zazie Todd, Alice Deignan & Lynne Cameron (eds.), *Human cognitive processing*, vol. 26, 209–244. Amsterdam: John Benjamins.

Zbikowski, Lawrence M. 1997. Conceptual models and cross-domain mapping: New perspectives on theories of music and hierarchy. *Journal of Music Theory* 41(2). 193–225.

Zbikowski, Lawrence M. 2008. Metaphor and music. In *The Cambridge handbook of metaphor and thought*, 502–524. Cambridge: Cambridge University Press.

Zbikowski, Lawrence M. 2018. Conceptual blending, creativity, and music. *Musicae Scientiae* 22(1). 6–23.