



OPEN ACCESS

EDITED BY

Vito Evola,
Universidade NOVA de
Lisboa, Portugal

REVIEWED BY

Isabella Poggi,
Roma Tre University, Italy
Katharina J. Rohlfing,
University of Paderborn, Germany

*CORRESPONDENCE

Katharina Meissl
katharina.meissl@kuleuven.be

SPECIALTY SECTION

This article was submitted to
Multimodality of Communication,
a section of the journal
Frontiers in Communication

RECEIVED 05 July 2022

ACCEPTED 06 October 2022

PUBLISHED 04 November 2022

CITATION

Meissl K, Sambre P and Feyaerts K
(2022) Mapping musical dynamics in
space. A qualitative analysis of
conductors' movements in orchestra
rehearsals. *Front. Commun.* 7:986733.
doi: 10.3389/fcomm.2022.986733

COPYRIGHT

© 2022 Meissl, Sambre and Feyaerts.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Mapping musical dynamics in space. A qualitative analysis of conductors' movements in orchestra rehearsals

Katharina Meissl^{1*}, Paul Sambre² and Kurt Feyaerts¹

¹Department of Linguistics, Faculty of Arts, University of Leuven, Leuven, Belgium, ²Department of Linguistics, Faculty of Arts, University of Leuven, Antwerp, Belgium

In this contribution, we examine the way in which orchestra conductors use the space around them to convey aspects of musical dynamics. In music, *dynamics* refers to the intensity of volume of notes and sounds and its interpretation is highly context-bound. We approach dynamics as a phenomenon of emergent and construed meaning in interaction, induced by the music score and the interpretation *in situ* by musicians and the conductor. Conductors' movement-based instructions on dynamics result in highly complex usage events. This study aims at disentangling these instances by asking how conductors move and use the space around them to instruct on (un)desired aspects of musical dynamics, zooming in on movement direction as a central formal feature. We find ourselves at the crossroads of cognitive and interactional linguistics, aligning with existing studies on the interactional and contextually embedded nature of music interaction. From a cognitive linguistic perspective, this endeavor translates as the identification of the construal mechanisms (metaphor, specificity and viewpoint) that underlie and therefore motivate movement directions in the specific instances under examination. The analysis is based on 10h of video data from a corpus recorded during rehearsals of five conductors instructing their respective orchestras in Dutch. Our data reveal that conductors use different movement patterns, some of which appear to involve opposite movement directions for expressing a similar music dynamical aspect, e.g., depending on the usage event, a vertical upward movement can mean both a request for playing louder and softer. By taking into account different construal mechanisms, we are able to provide an encompassing multimodal analysis, in which these allegedly deviating oppositional movements appear as consistently motivated (metaphorical) expressions, which profile a similar target concept involving different viewpoints.

KEYWORDS

musical dynamics, orchestra conducting, construal, metaphor, viewpoint, kinesemiotics

Introduction

In Western-European classical music, the role of the conductor is a quite unique one, given that conductors are the only performers who do not actively produce musical sound (Schuldt-Jensen, 2015, p. 386). This fact creates the paradox of conductors coordinating the musical performance without actually having its production in their own hands (Watson, 2012, p. 18). Their responsibilities are manifold: on a very basic level, they need to “direct the musical traffic” (Boyes Braem and Bräm, 2000, p. 146), thus indicating when (parts of) the orchestra should start and stop playing as well as beating time, which is conventionalized to a higher degree in specific beat patterns than other aspects of conducting. On a higher level, the conductor has to conceptualize the complete sounding realization of a musical piece as an interplay of the composer’s aspirations as noted in the score, their own interpretation of it as well as the musicians’ performance (Schuldt-Jensen, 2015, p. 386). Underlying this complex task is a broad set of expertise, ranging from musical knowledge to leadership and communicative skills (Watson, 2012, p. 18–19).

Among the many aspects that conductors express communicatively toward the orchestra is *dynamics*, which is the focus of this contribution. In music, the term *dynamics* relates to “the intensity of volume with which notes and sounds are expressed” and is “one of the fundamental parameters of composition which function interdependently to create musical meaning and structure” (Thiemel, 2001). Other parameters relevant for performance also conveyed by conductors include rhythm, tempo, phrasing, articulation, timbre and balance. These parameters can, but do not necessarily have to be, noted in the score.

In the score, different aspects of dynamics can be indicated through measures such as *piano*, *forte* and degrees thereof, e.g., *mezzo-piano* and *fortissimo*. Changes in volume can be indicated as gradual in nature, increasing (*crescendo*), decreasing (*diminuendo/decrescendo*), or abrupt as in the case of accent and *sforzato*. Terms referring to dynamics should always be regarded as relative, not as absolute (Gehrken, 2006, p. 56). While indications of (changes in) volume can be louder or softer in relation to others, their exact loudness is not determined. “Thus *dynamics*, perhaps even more than tempo, will be seen to depend on the taste of the performer or conductor” (Gehrken, 2006, p. 56).

This context-boundness of the interpretation of dynamics (Weeks, 1996, p. 248) is in line with a usage-based linguistic perspective on both processes of meaning making (Barlow and Kemmer, 2000) and musical performance. We view the conceptualization of musical dynamics as a phenomenon of emergent and construed meaning in interaction (Zima and Brône, 2015), induced both by the normative basis of the music score and the interpretation *in situ* by the conductor and musicians in an evolving joint practice. In a usage-based

model of language analysis, discursive practices associated with orchestral conducting emerge spontaneously and, in the context of a rehearsal, are dynamically updated by participants.

The current paper is structured as follows: In Section *Theoretical rationale and research aims*, we discuss the integration of multimodal interaction analysis of musical settings (Veronesi and Pasquandrea, 2014; Hsu et al., 2021) and cognitive linguistics (Dancygier, 2017) as well as kinesemiotic characteristics of conductors’ movements (Maiorani, 2020). This rationale results in the formulation of our research aim to identify construal mechanisms underlying movement direction patterns in situationally embedded instructions. Section *Materials and method* provides information about the video corpus and the method adopted for this contribution. The analysis of seven authentic examples follows in Section *Analysis*. In Section *Summary*, we summarize our analysis on movement direction patterns pertaining to dynamics, construal mechanisms and especially viewpoint phenomena. The findings are discussed in Section *Discussion*.

Theoretical rationale and research aims

With regard to mapping the kinesemiotic alignment of usage events in which conductors instruct their orchestra about the way in which certain phenomena in the domain of musical dynamics are (not) to be performed, our multimodal analysis is inspired by both an interactional and a cognitive approach. We highlight the complementarity of both paradigms as our analysis reveals the importance of situationally embedded and interactionally driven resources as well as the identification of cognitive construal mechanisms along whose lines the kinesemiotic alignment of the movements under scrutiny may be motivated.

Interactional studies on orchestra conducting

We align ourselves with existing studies that regard orchestral interaction as shaped significantly by its situatedness, being embedded in a very specific physical context and embodied in the sense that music-making and conducting rely on the use of the body (Parton, 2014, p. 405). A range of work from a conversation analytical and interactional linguistic perspective scrutinize interaction in different musical (instructional) settings (Szczepiek Reed et al., 2013; Tolins, 2013; Veronesi and Pasquandrea, 2014; Sambre and Feyaerts, 2017; Hsu et al., 2021; Ivaldi et al., 2021; Sambre, 2021), focusing on different aspects of embodiment and multimodality as well as the expression and negotiation of instructions. For orchestra rehearsals specifically, the communicative repertoire

of conductors and interaction in the rehearsal setting has been studied with regard to its sequential and multimodal organization (Meissl et al., Submitted; Weeks, 1985, 1996; Veronesi, 2014; Stoeckl and Messner, 2021), conducting gestures (Boyes Braem and Bräm, 2000; Parton and Edwards, 2009), facial expressions used by conductors (Poggi, 2002), vocalized and sung instructions (Messner, 2020) as well as the negotiation of epistemic stance (Parton, 2014).

Similar to other instructional settings, the rehearsal process aims at improving the collective performance of the orchestra, mostly working toward a concert. Rehearsals are institutionalized to a high degree and are thus shaped by a hierarchical distribution of participant roles, in which the conductor, as a default, has the right to speak, while musicians' turns, either to speak or to make music, are prompted either by the conductor or the musical score (Stoeckl and Messner, 2021, p. 2). The degree to which this hierarchy is adhered to, however, varies with each individual conductor-ensemble constellation (Schuldt-Jensen, 2015, p. 388). Next to the relatively stable participant roles, there is a common sequential order for rehearsals. They generally consist of alternating sequences of play and interruption of play. The former are marked by music-making and conducting or so-called instructions "on the fly" by the conductor, while the latter usually contain instructions and discussions of the music just played or to be played next (Messner, 2020, p. 318). Within these larger sequences, several activities on different levels take place, including evaluating, locating, demonstrating, clarifying, describing and signaling (Stoeckl and Messner, 2021, p. 6). The coordination of different activities linked to conducting as well as the progression of the rehearsal combined with the one-to-many constellation makes orchestra conducting a multi-activity (Haddington et al., 2014) par excellence. Conductors need to constantly shift their focus of attention to different aspects of performance and also physically re-orient toward different participants. This results in the gesture space of the conductor shifting constantly due to the affordances (Gibson, 1979) of the physical constellation and the tasks at hand.

For the communication with the orchestra, conductors use spoken or sung language as well as their whole body, drawing upon manual gesture, torso and head movement, body posture, facial expressions and gaze, and sometimes also movement of the legs. Traditionally, there is a division of labor between each of the conductor's hands, according to which the right, or dominant, one is responsible for the beforementioned coordination of tempo and rhythm, while the left, or non-dominant, hand is used for adding information about sound colors, musical phrasings or dynamics (Boyes Braem and Bräm, 2000, p. 245). In this division, the movement of the right hand is more conventionalized, while the use of the left hand appears to be used more for expressing individual interpretation. However, the degree to which this tradition is adhered to is highly dependent on individual conductors. A large part of the art of conducting

remains, as Watson (2012, p. 22) puts it, intangible and is often attributed to "personality, charisma and power of persuasion."

However, certain patterns that shape a conductor's communicative repertoire are identifiable and have been studied from different perspectives. Spoken as well as sung or vocalized instructions constitute an important part of rehearsal work (Weeks, 1996; Messner, 2020), but they often co-occur with other semiotic resources. For example, Stoeckl and Messner (2021, p. 12) identify speech, gesture and gaze in combination as frequent and generic in conductors' instructions, often complemented seamlessly by vocalizations when musical passages or sound qualities are depicted (Clark, 2016). Gaze as well as body position and movement are, according to Stoeckl and Messner (2021, p. 12), primarily used for addressing and orienting¹. Facial expressions in conducting have been found to be not merely idiosyncratic but systematic in use by Poggi (2002) who describes "the lexicon of a conductor's face." The notion to create grammars or vocabularies of conducting movement is also present in a study by Boyes Braem and Bräm (2000) who attribute metaphorical meanings to certain types of gestures used by conductors. Some of the gestures described in that study, however, also occur in the data set of this contribution with different meanings and nuances. For example, a conductor touching or pointing at their own ear is described as asking for correct intonation by Boyes Braem and Bräm (2000, p. 159), but can also refer to *balance*, an aspect of dynamics, as surfaces in the data set we study. This leads us to conclude that the communicational repertoire of conductors is not that fixed.

Conducting instructions pertaining to dynamics have been studied both from a more quantitative approach by Opazo (2018) as well as qualitatively by Poggi and Ansani (2016) and Poggi (2017). In these studies, video data were analyzed focusing on specific aspects of musical dynamics: *piano*, *forte*, *crescendo*, *diminuendo*, and in the case of Opazo (2018) also the more fine-grained *pianissimo* and *fortissimo*. For instructions on these aspects, a combination of different characteristics was studied: formal features of manual gestures (e.g., handshape, orientation, location), movement parameters (direction, velocity, duration) as well as parameters of expressivity (amplitude, tension, fluidity). Next to certain common handshapes and orientations, such as the fist for loud sounds or a flat hand palm down for soft sounds, tension has been shown to be higher for louder and lower for instructions on softer sounds (Poggi, 2017, p. 41–42; Opazo, 2018, p. 110–111). Movement amplitude is another important parameter for sound volume, with evidence that the larger the movement amplitude, the higher the sound volume that is expressed (Poggi, 2017, p. 41, 43). Similar to the amplitude

¹ However, our forthcoming study on so-called "contrast pairs" (Weeks, 1996) suggests that body movement does indeed play an important role in the expression of instructions about music and sound quality, similar to facial expressions (Meissl et al., Submitted).

of conducting gestures, upward movement to express loud(er) sounds and downward movement for soft(er) sounds have been described in the studies listed above, see also Section *Meaning construal*. However, an aspect which, to our knowledge, has not been studied systematically so far concerns the question whether conductors' dynamics-oriented movements reveal any co-occurrence patterns, along the lines of which certain aspects of musical dynamics are preferably expressed by a specific movement direction on a spatial axis (vertical, horizontal or sagittal). In case direction patterns should emerge, it will be investigated to what extent they can be motivated by situated phenomena of the interaction and/or underlying cognitive construal mechanisms.

Therefore, before formulating our research aims and embarking on the analysis of our data, we take a step back from the specific musical setting in order to familiarize ourselves with more general concepts proposed in cognitive linguistics regarding the construal of meaning.

Meaning construal

Beyond the interactional linguistic perspective, through which we integrate local and situational resources in the analysis, the present study also adopts a cognitive linguistic perspective on the process of musical meaning making (Cox, 2016; Zbikowski, 2017; Spitzer, 2018; Antović, 2019). Choosing a usage-based approach, we view interaction as the integration of all available semiotic resources used for linguistic meaning-making including bodily, visual and acoustic features as they unfold over time (Langacker, 2008, p. 73; Langacker, 2010, p. 90–95; Ladewig, 2020, p. 179), thus extending the phonological pole of a construct (Langacker, 2008, p. 457).

Casad (1995, p. 23) rightly points out that an interactant's "ability to conceptualize situations in a variety of ways is, in fact, the foundation of cognitive semantics." To achieve their communicative purpose, interlocutors have a wide range of so-called construal mechanisms at their disposal. Several typologies of construal operations² have already been proposed, e.g., by Langacker (1987, 1991), Taylor (2002), Talmy (2003). Croft and Cruse (2004, p. 45) present an overview of the relevant literature on construal and propose a typology of construal operations, grouped along the general categories of attention and salience, comparison, perspective and viewpoint, and Gestalt (Feyaerts, 2013, p. 207–209). Along the lines of these construal operations, language users may decide to share their experiences in a variety of ways, for instance by conceptualizing the lower parts of a mountain in terms of its *foot*, or by referring to the change of seasons in terms of travel-like movements as in *Spring is approaching fast* or

Finally we can leave Winter behind, etc. Taking a multimodal perspective on interaction, these construal mechanisms can surface not only in speech but also in other semiotic resources (Cienki, 2022).

For our present purpose, we focus our attention on the impact of metaphor, specificity and viewpoint as the most prominent construal mechanisms underlying the expression of (un)desired realizations of musical dynamics.

Firstly, we consider *metaphor*, whose ubiquity has been abundantly described in studies situated within the framework of Conceptual Metaphor Theory (CMT, Lakoff and Johnson, 1980; Lakoff, 1987; Kövecses, 2015; Kok and Cienki, 2016). CMT defines metaphors as systematic cognitive mappings between two conceptual domains. One of these domains, which is complex but not necessarily abstract is called the target. The inner logic of the target domain is represented in terms of the inner logic of another domain, the source, based on some sort of similarity relation. Such metaphorical mappings structure our experience as well as the communication about our experiences (Prové and Feyaerts, 2022). In the present study, we investigate by what imagery aspects of musical dynamics are construed metaphorically in terms of movement in 3D space. In line with previous studies of the spatial hence metaphorical mapping of the concept of musical pitch in terms of verbal expressions of verticality, according to which we speak of *high* and *low* tones, *climbing* and *falling* arpeggios, etc. (Zbikowski, 2002; Shayan et al., 2011; Cox, 2016; Prové and Feyaerts, 2022, among others) we expect to find co-occurring gestural resources for the metaphorical expression of musical dynamics as well. As observed by Eitan (2013, p. 173–176) the domain of verticality or height does not only serve the metaphorical structuring of pitch, as also loudness relations appear to be categorized along the same lines of vertical logic. The mapping of sound volume as size, thus, LOUDER IS BIGGER and SOFTER IS SMALLER as well as the vertical mapping of LOUDER IS UP and SOFTER IS DOWN have already been identified as highly salient in conductors' instructions on musical dynamics (Opazo, 2018, p. 113–114). What qualifies these metaphorical structures as 'patterns' is the systematic, non-arbitrary mapping of the target concept onto a specific alignment of the source concept. Hence, the concept of increasing volume cannot be represented by just any concept in the domain of verticality. Instead, only expressions referring to tall or big objects (*make this a huge forte!*) or upward processes (*climb to that forte!*) can be used to profile increasing loudness. Considering our focus on sound volume and intensity, the current study pays specific attention to metaphors expressed in specific movement patterns used by conductors, according to which, for instance, increasing loudness can be expected to be represented visually by an increasing, growing gestural movement situated on (a combination of) the vertical, horizontal and sagittal axis. Such metaphorical mappings may be grounded in spatial gestalts and force dynamics, relating to both natural or humanly exerted

² We use the terms of construal "operations" and "mechanisms" interchangeably.

force (Talmy, 1988)³. Larson (2012, p. 23, 329) demonstrates that notions such as gravity, attraction or other aspects of physical force structure the experience of musical concepts such as melody, meter, rhythm, and tempo in terms of movement in space (Feldman et al., 1992; Johnson and Larson, 2003, p. 75). On a gestural level, metaphorical mapping onto movement in space surfaces in hand or movement shapes revealing SOURCE-PATH-GOAL, CONTAINMENT or other spatial gestalts (Mittelberg, 2018). Characteristics of conceptual PATHS, such as their *directionality*, serve as observable structures which in language and discourse act as interface between sensorimotor experience and conceptualization (Johnson, 2017, p. 86). Instructional conducting movements clearly exploit the spatiotemporal experiential basis of force and movement in space, as, for example, with a downward hitting gesture for a hard sound quality (Boyes Braem and Bräm, 2000, p. 154).

Secondly, the construal operation of *specificity* is omnipresent as it pertains to the level of granularity at which we conceptualize and communicate our experiences (Cienki, 2022, p. 4). For various communicative reasons (expressivity, euphemism, accuracy, humor, etc.) we may decide to use more (or less) detailed descriptions to refer to objects, properties, processes, etc. or to express a stance toward them. In a medical situation, for instance, depending on situational factors like age, previously shared knowledge, relationship and emotional state between the patient and the interlocutor, a doctor may decide to refer to a patient's deadly disease in more general, euphemistic terms like *autoimmune* or *lingering* or *chronical disease*, rather than using a more specific terminology like *lung cancer*. Conversely, along the lines of the same construal mechanism, in their report about the result of a championship's race, a journalist is expected to communicate which medal a race favorite has won, a *golden*, *silver* or *bronze medal*. However, in their interview, the athlete finishing third may express their joy about having won a *medal* without specifying its color. With respect to the overarching phenomenon of *musical dynamics* as the topic of the present study, more specifically, partially overlapping categorizations like *loudness*, *intensity*, *diminuendo*, *crescendo*, *accent*, etc. may be used for different communicative reasons and at the same time trigger different verbal and gestural metaphors.

³ Closely related to metaphorical mappings are so-called *force dynamics* (Talmy, 1988), another construal mechanism that allows us to conceptualize abstract and complex knowledge structures. Force dynamics is a system to express experiences of force exertion, opposition or resistance and overcoming (Talmy, 2003, p. 232) and similarly to image schemas can surface on different levels of metaphorical reasoning (Mittelberg, 2017). These embodied schemas derive directly from our non-mediated bodily experience with the world and organize our experience and comprehension (Johnson, 1987, 29).

As a third construal mechanism, our analysis takes *viewpoint* into account, which pertains to the inherent perspective through which any conceptualization is determined (Sweetser, 2012; Cienki, 2022). In line with a socio-cognitive account of meaning, we understand viewpoint in terms of one's personal perspective on a certain issue, expressed in an intersubjective stance-taking act, which emerges in a constant coordination process of perspective-taking and mentalizing among different interlocutors (Feyaerts et al., 2017). With this view, we side with the Theory of Mind (Whiten, 1991; Tomasello, 1999; Givón, 2005), which identifies our ability to conceptualize thoughts, ideas, emotions, attitudes, beliefs, etc. in other people's mind as a unique human capacity (Brône, 2010, p. 91–92). While interacting, participants imagine what they assume to be in the minds of their conversational partners aligning their construal with it so that, ultimately, conversation emerges as a “process that requires constant alignment and negotiation among *intersubjective viewpoints*” (Feyaerts and Oben, 2014, p. 277–278; Verhagen, 2015, p. 238–240). As already mentioned with regard to the application of metaphor, language users also get to decide from which viewpoint an action or situation will be communicated. “Even if viewpointing is not consciously intentional, linguistic constructions are infused with viewpoint [...] to the point where these are conventionalized” for speakers/signers and addressees so that the latter make inferences about the viewpoint of the former, resulting in “joint construal intersubjectively” (Janzen, 2022, p. 6; referring to Traugott and Dasher, 2001). Locating Mary's house, for instance, may be formulated using an external viewpoint as in *Mary lives on the left bank of the river* or by using an internal viewpoint as in *Mary lives across the river*, etc. The choice of the latter variant nicely illustrates the intersubjective dimension of operating viewpoint as a construal mechanism, as it subsumes that the producer of the utterance assumes their interlocutor(s) awareness of the producer's location vis-à-vis the river.

Related to viewpoint, Sweetser (2012, p. 1) highlights the relevance of embodied experience as well as spatiality in reference to gesturing bodies (Sweetser and Sizemore, 2008). Also in Kinesemiotics (Maiorani, 2020), the interplay between the human body and space for the process of meaning-making is foregrounded. In this paradigm, the relation between body and space is regarded as a dual one: physically, a body is located relative to other bodies and it is subject to physical laws of nature. Contextually, or culturally, “a body occupies space as a semiotic dimension, a three-dimensional map of meaningful areas” (Maiorani, 2020, p. 26–27). Not movement of the body as such is inherently meaningful, but rather the projections it makes onto space relative to other participants or objects. In the case of conductors, they are both physically and culturally placed in the center of attention, immediately rendering the space between them and the musicians semiotically charged.

Research aims

At this point, we are well-positioned to formulate the main research aims underlying the present study. At the most general level, we study the question how conductors move and use the space around them to instruct on (un)desired aspects of musical dynamics. In order to highlight the salience of the interaction between moving bodies and 3D space, we isolate movement direction as a formal parameter to identify patterns in conductors' instructions, along the three major dimensional axes: verticality, horizontality and sagittality. In line with existing research on verbal metaphorical patterns in the domain of musical experience (see Section *Meaning construal*), the present study zooms in on co-occurrence patterns, in which certain aspects of musical dynamics are represented by specific movement directions along these spatial axes. Along with the identification of these co-occurrences, we will investigate what motivates them, cognitively and interactionally.

On the level of a cognitive linguistic analysis of the selected usage events, this endeavor translates as the description of the construal mechanisms that underlie directional patterns and systematically surface in the instances under concern. Lastly, we aim to shed light on the benefits of enriching studies of human face-to-face interaction by taking into account metaphor, specificity and viewpoint phenomena in relation to both the spatial arrangement of participants and the object of conceptualization, in this case, musical dynamics. With this research focus, we side with Cienki (2022, p. 12), advocating the study of movement and gesture as “an inherently spatial medium of expression, [which] can allow future empirical research in cognitive science [...] concerning spatial cognition as being a fundamental basis for how we conceptualize more abstract domains.”

Materials and methods

We use video recordings of five different conductors during rehearsal with their respective wind and brass orchestras in Flanders, Belgium (Simon and Feyaerts, 2020). The conductors did not receive any specific instructions other than to go about the rehearsals as they usually would with the amateur or youth ensembles. In the course of several weeks, three rehearsals per conductor were filmed resulting in about 30 h of material. During recording, the camera was placed behind the musicians in a way that only the conductor is fully visible in the video frame. An additional microphone was placed on the stand in front of the conductor to ensure proper recording of verbal instructions and vocalizations; the language used in the corpus is Dutch. For this contribution, a sub-corpus of one rehearsal per

conductor was delimited for reasons of feasibility, amounting to about 10 h of data.

To identify audible and visual instructions pertaining to musical dynamics, we used the annotation software ELAN (Wittenburg et al., 2006). We included the following usage events in this process, regardless of whether they are produced while the orchestra is playing or while playing is interrupted: (1) verbal utterances referring to sound volume and intensity or changes in volume, e.g., “crescendo,” “really fortissimo and then back,” “here it's far too loud,” (2) vocalizations and singing in proximity to either a verbal utterance referring to dynamics or accompanied by movement and (3) movement-based communication pertaining to volume and intensity (a) along-side verbal utterances, (b) along-side vocalizations and singing, or by (c) movement only.

For movements to be taken into account, they have to noticeably deviate from the regular beating of time in their close sequential surroundings. As soon as the beating of time is temporarily either accompanied or replaced by different movements with any body part which seem to refer to an aspect of musical dynamics, this movement was noted. When beating time remains consistent in shape but changes significantly in amplitude, this was noted as well. Importantly, the ‘default’ conducting movement is an idiosyncratic value which varies significantly across conductors and largely depends on the different musical pieces being performed and was therefore established on an individual basis.

Four aspects were considered to verify whether instructions refer (primarily) to musical dynamics. First, we checked for verbal utterances which explicitly refer to dynamics during or sequentially close to movements. Secondly, non-lexical vocalizations such as singing or ‘shushing’ at certain acoustic volumes aided the disambiguation. Thirdly, similar to the next-turn proof procedure in conversation analysis (Hutchby and Wooffitt, 2008, p. 13), we took into account how musicians audibly adjust their performance in response to the instructions at hand. The notion of an embodied next turn proof procedure, where understanding is signaled not only after, but already during a turn, as discussed by Goodwin and Salomon (2019, p. 5) was useful to interactionally frame what happens during conducted orchestra performance, in which instruction and performance occur almost simultaneously to each other. Lastly, highly conventionalized gestural movements were categorized accordingly, as when an extended index finger held in front of the mouth signals silence. In other words, the segmentation of instructions pertaining to dynamics, following a usage-based approach, was based on visual and acoustic cues in the video data and not on the musical score. We decided to focus on what the conductors make relevant in interaction through multimodal practices (Mondada, 2019a) in order not to miss cases, either, where there are

indications of dynamics in the score but the conductor does not convey them through movement, or, where there are no specific indications noted but the conductor does perform relevant instructions.

This segmentation of instructions pertaining to musical dynamics resulted in roughly 1,100 units, which served as a first overview and allowed us to get acquainted with the variety of the cases under scrutiny. In these segments, when present, speech was transcribed based on the concept of intonation units (Chafe, 1994, p. 93). However, as the focus for this contribution lies on movement-based communication, verbal instructions lacking any other semiotic resources were not included in the closer selection.

To analyze movement-based instructions, we chose “Gesture Units” as the unit of analysis, spanning “from the moment the hands leave rest until the hands return to rest” (Rohrer et al., 2020, p. 13), where “rest” is to be regarded as relative and can differ in the degree of relaxation of hand and/or arms depending on different factors such individual styles, handling of objects or environmental conditions. For this contribution, possible rest positions may include interlacing the hand in front of the body, or having them at the side of the body or also the default of beating of time which in essence does not pertain to musical dynamics.

Although conducting movement cannot be fully equated with co-speech gesture, regarding conducting as gesture units and the division into movement phases along with the annotation of formal gestural features facilitate an accurate analysis of complex movements (Ladewig and Bressemer, 2013; Rohrer et al., 2020). Importantly, we do not aim at performing a formally strict gesture analysis. Rather, we adopt certain practices to aid our study on movement directional patterns. Consequently, conducting movement was divided into gesture phases, in order to identify salient parts of the previously segmented units. For strokes, movement directions were noted along the vertical, sagittal and horizontal axis, which served the identification of movement patterns linked to musical dynamics, according to common practice in metaphor analysis (Cienki, 2017). Such patterns are, for example, vertically upward and/or sagittally forward movement to express louder sounds, and thus, a value on one or several spatial axes corresponding to a specific meaning. Several patterns of that kind will be discussed in Section *Analysis*. For this contribution, we considered (combinations) of movement directions as patterns when they occurred across all five conductors in our corpus.

In light of the fact that visual access to the conductors’ actions is confined to one single camera perspective, which is behind the orchestra, right opposite to the conductor’s forward-looking position, determining a movement’s alignment along one or more of the spatial axes requires special attention in terms of assuming a flexible analytical perspective along the conductor’s constantly shifting orientation toward individuals or (sub)groups across the orchestra.

Analysis

In the following sections, we describe our findings guided by seven authentic corpus examples⁴, which were chosen as prototypical instances of the patterns emerging in our data. Rather than structuring our analysis along the different construal mechanisms, we depart from the spatial dimension with regard to movement direction. In Section *Mapping the prototype: Increasing intensity as expanding size*, we scrutinize patterns that have been identified as predominant in previous literature. We enrich these findings by focusing on the interplay of metaphor, specificity and viewpoint. In Section *Complicating the picture: Specificity and viewpoint*, we show alternative ways for construing aspects of musical dynamics in our data and how they can be motivated both cognitively and interactionally. Each example will be examined along the following lines: first, we provide an observational description of the video fragments, focusing on movement directions, which is then followed by a discussion of the construal mechanisms mentioned above.

Mapping the prototype: Increasing intensity as expanding size

Previous work has uncovered several conceptual metaphors serving the purpose of representing the acoustic experience of music and sound in general, one of the most predominant ones being the mapping of VERTICALITY or SIZE onto LOUDNESS. These are highly schematic, so-called generalized metaphors, which can be elaborated in more specific metaphorical imagery like *climbing a ladder* or *producing a big, overwhelming sound*, respectively. In the specific setting of conducting (see Section *Interactional studies on orchestra conducting*), a general observation concerns the amplitude of the beating of time performed by a conductor. Amplitude can be an indicator of the preferred sound intensity at that moment, as described by Watson (2019) and which is also reflected in conducting manuals (for a summary, see Sousa, 1988, p. 34). Indeed, general sound volume as well as a more global atmosphere of (parts of) musical pieces can often be deduced from the size of the movements conductors make. Beating of time with a small amplitude is likely to suggest a low sound volume while a larger amplitude of motion can suggest more intensity. These observations have been confirmed by Poggi (2017) and Opazo (2018) and they also hold for our data set.

⁴ The transcripts below are based on the conventions for transcribing multimodal data as suggested by Mondada (2019b). Arrows were added onto stills when considered an added value. Conductors’ manual gestures are marked with *, body and head movements with \$. Actions from the orchestra are marked with †.

However, amplitude alone is by far not the only possible resource to indicate aspects of dynamics. In order to be able to zoom in on the ways in which conductors exploit the space around them to conceptualize their interpretation of the music played, we will further explore the spatial alignment of these movements along the vertical, horizontal and/or sagittal axis including their directionality on each of these axes.

Along the vertical axis, we see a pattern of loud(er) sounds being depicted as higher up than soft(er) sounds, based on the metaphor LOUDER IS UP and SOFTER IS DOWN. The excerpt in Figure 1A shows this tendency quite clearly⁵.

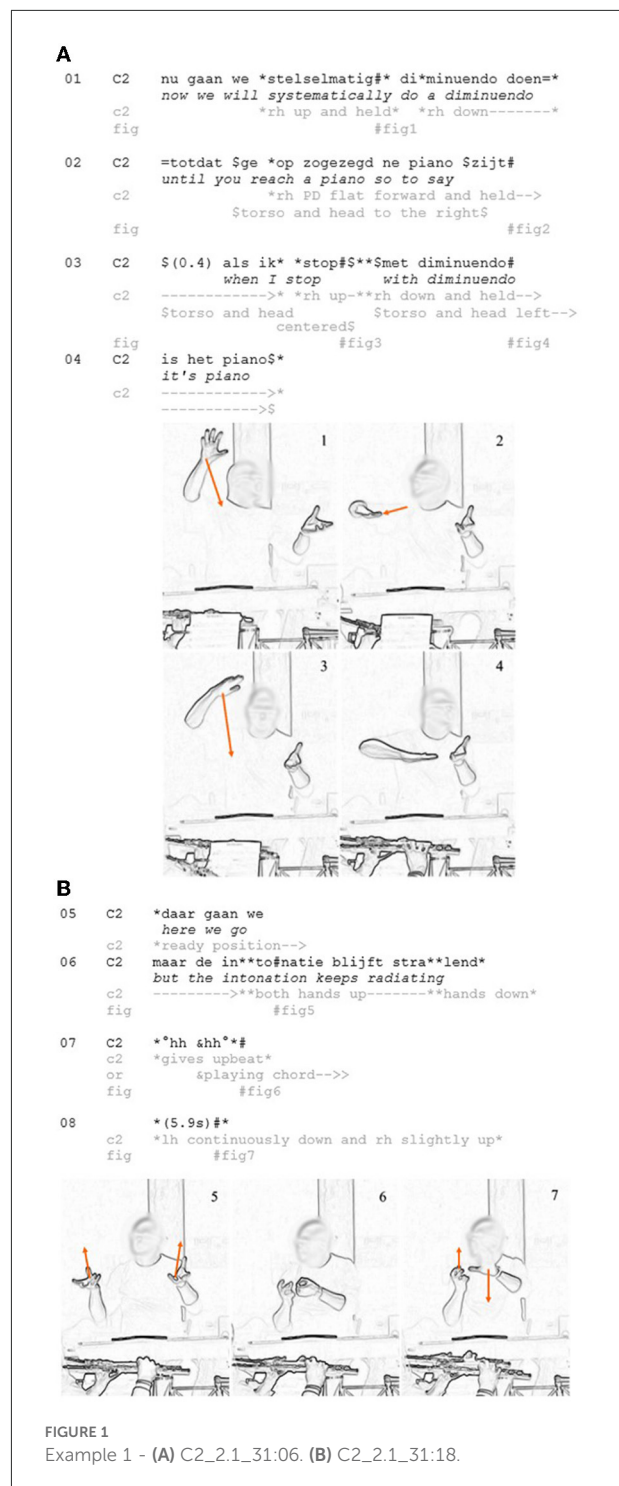
Just before this example, the conductor has been working on intonation across different instrument sections in order to reach clean relative pitches. He has taken up a ready-position for conducting with both hands, signaling that he is about to start another playing sequence. However, before starting, he instructs the musicians to continuously decrease the sound volume (line 01), while they play the respective notes. The conductor raises his right hand over head height in preparation for the stroke of the gesture, a continuous downward movement along a vertical axis, representing a decrease in sound volume (*diminuendo*). After this downward movement, his flat hand is facing palm down and he moves it as if flattening a surface, that is the volume to be reached, namely a piano (line 02). Directly after, he raises his hand to head height again to repeat the downward movement, this time more quickly and holding his hand in the final position at chest height (line 03).

In this example, louder sounds are conceptualized as vertically higher up than softer sounds, which corresponds to findings in earlier studies on orchestra conducting such as Boyes Braem and Bräm (2000, p. 159) and maps onto the VERTICALITY metaphor. The combination of the palm-down orientation with a flat hand and downward movement to signal softer sounds, used by the conductor in this example, has also been observed by Poggi (2017, p. 43–44) and Opazo (2018, p. 79–80).

Additionally, the movement in this excerpt qualifies as the instantiation of the SOURCE-PATH-GOAL metaphor, involving a clear end point, marked by the movement that we described above as flattening a surface. This gesture, as well as the hold after the second and quicker downward movement the conductor makes with his right hand, can be regarded as delimiting the desired scope of intensity. The conductor's left hand is kept in place, still at the originally assumed ready-position and thus serving as a reference point in the gesture space.

Regarding viewpoint, we see that, although the use of different personal pronouns (*we*, *you*, *I*, *it*) marks lexical viewpoint switches, the performed gesture does not impact the perspectivized relation among the participants. Throughout the

⁵ The video clips of all examples can be found in the Supplementary material.



excerpt, the imagined sound quality is depicted from an external perspective, rather than from an internal one, which could be either the conductor's or the musicians' perspective.

As mentioned in Section *Meaning construal*, VERTICALITY can also serve as source domain for the concept of pitch. To illustrate how verticality may serve the metaphorical structuring

of dynamics and pitch as two different target concepts, even simultaneously, we analyze Figure 1B, which contains the playing sequence following the instruction on *diminuendo* described in Figure 1A.

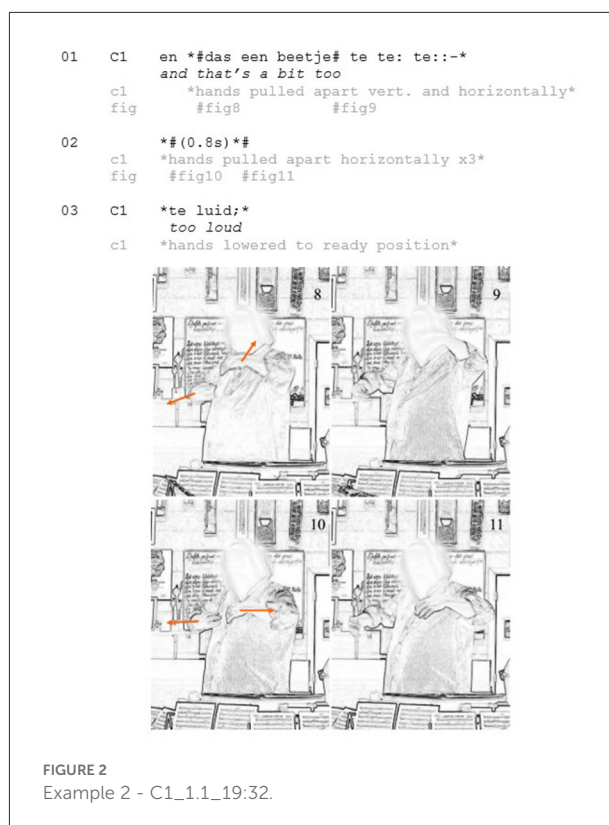
After the excerpt in Figure 1A, the conductor reassumes a ready position, but decides to add another instruction, this time referring to pitch. He says the intonation should “keep radiating” (line 02) and lifts both of his hands to head height facing palm up. At the same time, the conductor tilts his head back and directs his gaze up while raising his eyebrows. He initiates the playing sequence by giving the upbeat, and then holds both of his hands at a central position in his gesture space while the musicians play one single chord. The conductor lifts his left hand up to head height and extends it toward the musicians, the palm facing diagonally toward them and down. He moves his left hand downwards in a continuous motion, thus depicting the systematic *diminuendo* described in Figure 1A. As soon as his left hand has reached about half of the downward path, he raises his right hand facing the other way to chin height, again slightly tilting his head back and raising his eyebrows. This second action refers to the radiating sound quality the conductor asked for and has also been identified as expressing a “radiating” sound quality by Boyes Braem and Bräm (2000, p. 159).

In this example, the two target domains of dynamics and pitch are simultaneously conceptualized along a vertical axis, the depiction of each being attributed to a different hand, thus displaying a complex multi-activity of both monitoring and instructing on different performance aspects.

The complexity of this example also reveals a layering of different viewpoints when looking at both the form of the conductor’s right hand gesture and his facial expression. While the gesture is related to the imagined sound quality, the facial expression depicts the way in which an instrument has to be played in order to reach a higher pitch or rising intonation⁶. Thus, the internal viewpoint of the conductor conceptualizing the sound in front of his body is blended with the assumed viewpoint of the addressed musicians who need to produce a slightly higher pitched note in a single integrated scene. Most interestingly, both aspects and viewpoints are metaphorically motivated by essentially the same directionality on the vertical axis.

Like verticality serving as the source domain for different concepts in musical interaction, dynamics as a target domain is also metaphorically structured along the lines of other spatial dimensions, which are still very much in line with the LOUDNESS IS SIZE metaphor. The movements in example 2 critically evolve around the horizontal axis.

6 Importantly, the ensembles in the corpus consist of wind and brass instruments for the most part. In this excerpt, the conductor addresses musicians who produce sound with their instrument using air flow through their mouths.



Before the excerpt in Figure 2 starts, the conductor is working on intonation and has asked individual musicians one after the other to play a single note and to do so very softly. In the example, one musician plays the note in question, but significantly louder than instructed. The conductor comments on this performance by saying “and this is a bit too too too” (line 01) and gestures while trying to retrieve the right word to finish his utterance. He repeatedly pulls both of his arms outward from his body, the palms facing each other. After a short pause in his speech and three more repetitions of the outward movement he says “too loud” (line 03).

What we observe here, suggested by the orientation of the palms facing each other, is the basic conceptualization of a sound as an OBJECT or CONTAINER (Mittelberg, 2018, p. 12) expanding in size with increasing intensity in reference to the conductor’s body and gesture space, in which the center marks softer sounds and movement toward the periphery marks louder sounds. The gesture in Figure 2 is therefore also linked to the LOUDNESS IS SIZE metaphor and corresponds to the conceptualization of sounds as “thick” or “thin”⁷.

7 It should be noted here, that when referring to thickness of notes, there is often overlap with rhythm and thus the length of notes. Pitch is also conceptualized as thick and thin in some languages, e.g. Farsi (Dolscheid et al., 2013).

Regarding viewpoint, in this example there seems to be a clear involvement of the conductor's own body, from which the imagined sounds travel outwards, hence suggesting an internal, participant's viewpoint, in contrast to the stable external viewpoint described for Figure 1A. It may not be entirely clear whether this scene expresses the conductor mirroring the sound as it travels away from the musician producing it, thus assuming the latter's viewpoint, or, alternatively, the sound as experienced by the conductor himself, from his own perspective, independent of the musician. However, given that the conductor is offering an evaluation of the quality of the note just played, we may assume that it is indeed the musicians' viewpoint from which the sound is mapped in space.

While in the second part of the example in Figure 2, the movement is primarily performed on the horizontal axis, the first few repetitions of the expanding motion also involve some verticality, as the sound is being conceptualized as growing in height, as well as expanding away from the conductor's body sagittally, which leads us to the third spatial axis as yet another dimension of the LOUDNESS IS SIZE metaphor.

In the excerpt in Figure 3, the conductor is talking about the last note in a fragment that has been dealt with just before in the rehearsal. He asked the musicians to play an accent on each beat, which, however, resulted in them also performing an accent on the last note of the passage. Since this is not what the conductor had instructed them to do, he clarifies his request by first depicting the desired version (line 02). To that end, he vocalizes the last five notes of the fragment in question with a decrease in volume while simultaneously pulling his right hand toward his right shoulder. In doing so, the decreasing volume is mapped along a sagittal axis toward the conductor's body. In the second part of the sequence, he offers a depiction of the faulted version as previously played by the musicians (line 04). This time, the conductor vocalizes the same five notes, however, he increases the volume, ending quite loudly on the last note in relation to the previous depiction. Also, toward the last note of the vocalization, he pushes his right hand away from his body in a more intense forward motion while forming a fist⁸.

The underlying construal mechanism can in part still be linked to the LOUDNESS IS SIZE metaphor in the sense that louder sounds are construed as further away from the conductor's body than softer sounds. However, a strict interpretation of this metaphor does not suffice to account for

8 The fist as a handshape has been described by Bressemer and Müller (2014; p. 1584) as a recurrent gesture for German, with as its semantic core "strength, force and power" and an assertive quality, used for example to emphasize parts of utterances or signal emotional involvement. Analogously, musical dynamics as force, especially related to louder sound volume and higher intensity is mirrored in the use of the fist in Figure 3. As mentioned in Section *Interactional studies on orchestra conducting*, the fist has been shown to be a typical handshape for *forte* by Poggi (2017, p. 41) and Opazo (2018, p. 67).

the movement pattern at hand. The production of a louder sound by the musicians is depicted by the conductor with a forward motion away from his body, as if depicting the sound traveling away from its source. However, the retracting motion toward his own body cannot be interpreted based on the same logic. A softer sound, based on the metaphor of LOUDNESS IS SIZE, would have to be expressed in the same direction but with a smaller amplitude. However, this is not the construal operated by the conductor, which, rather, is a construal *ex-negativo*. The motivation of the retracting movement to depict playing softer derives only from its being opposed to the well-established metaphorical mapping of forward-motion expressing increasing loudness.

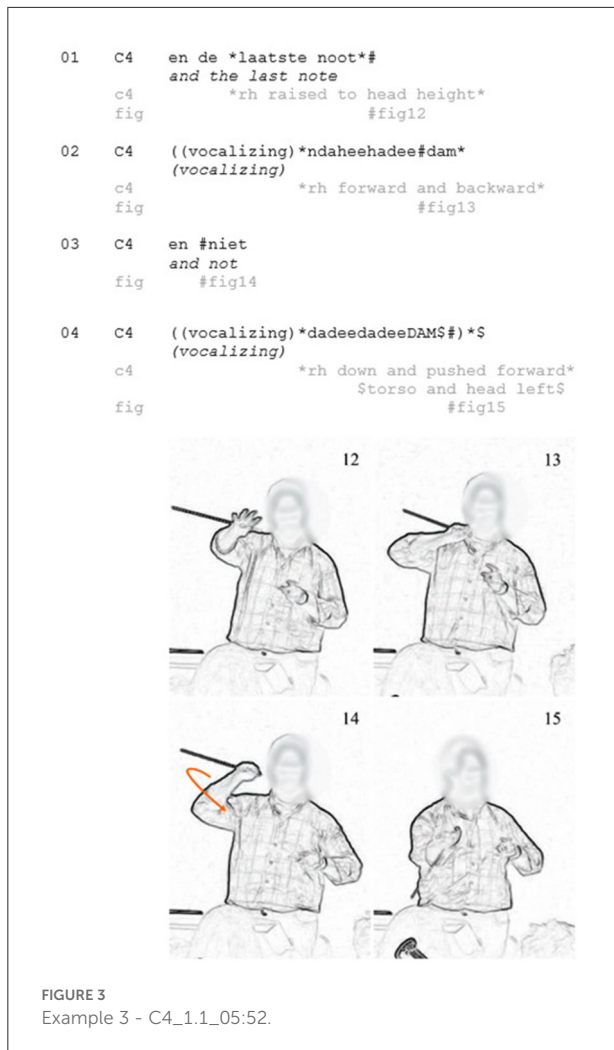
Whereas, in the previous examples, we witnessed a straightforward metaphorical construal of loudness along the lines of more or less isolated spatial dimensions, the present example reveals a much more complex construal involving not only a subtle interplay between various spatial axes, but also a shift from the conductor's viewpoint to the musicians', most clearly marked by the forward moving fist gesture⁹. This forward movement instantiates the construal of a force being exerted to project an objectified sound out of one's body. Crucially this movement only makes sense by taking into account a construed viewpoint switch, in which the conductor adopts the musicians' perspective of sound being produced by the interplay between their body and the wind instrument.

The analysis of the example in Figure 3 feeds into the critical understanding that the typical construal (and understanding) of musical dynamics (represented here as LOUDNESS) may not be restricted to the isolated metaphorical construal of that target concept in terms of the spatial logic as it occurs along either the vertical, horizontal and sagittal axis. It appears, instead, that the spatial mapping of various aspects of LOUDNESS in conductors' movements typically involve more than one or even all dimensions of spatial orientation. This finding is in line with the observation made by Schuldt-Jensen (2015, p. 395), who notes that the three-dimensionality of conducting movements has received little focus so far in teaching materials for aspiring conductors. Thus, scrutinizing the combination of movement directions seems crucial for an accurate analysis of both conceptual and interactional aspects of conducting.

The following two examples, expressing an intense *crescendo* and a *diminuendo*, respectively, require all three directional aspects in analyzing their metaphorical construal of musical dynamics. Notably, while the instructions in Figures 1–3 were given during an interruption of play, Figures 4, 5 are given on the fly, as the orchestra is performing.

In the excerpt in Figure 4, the conductor is already beating time at a relatively big amplitude and the general sound volume is quite high. Just before the music reaches a climax, he stands

9 Movement toward the musicians is also noted by Poggi (2017, p. 39) as an indication to play *forte*.

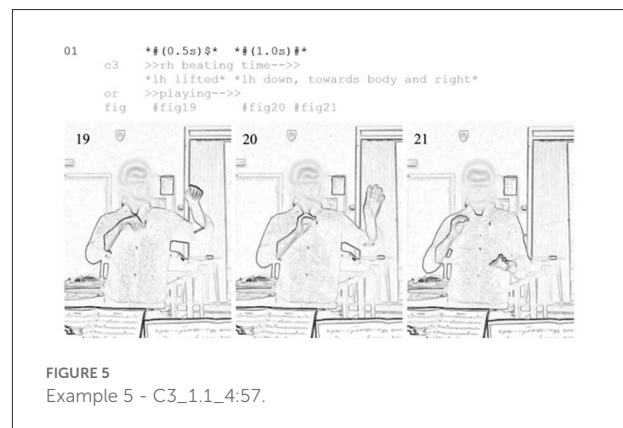
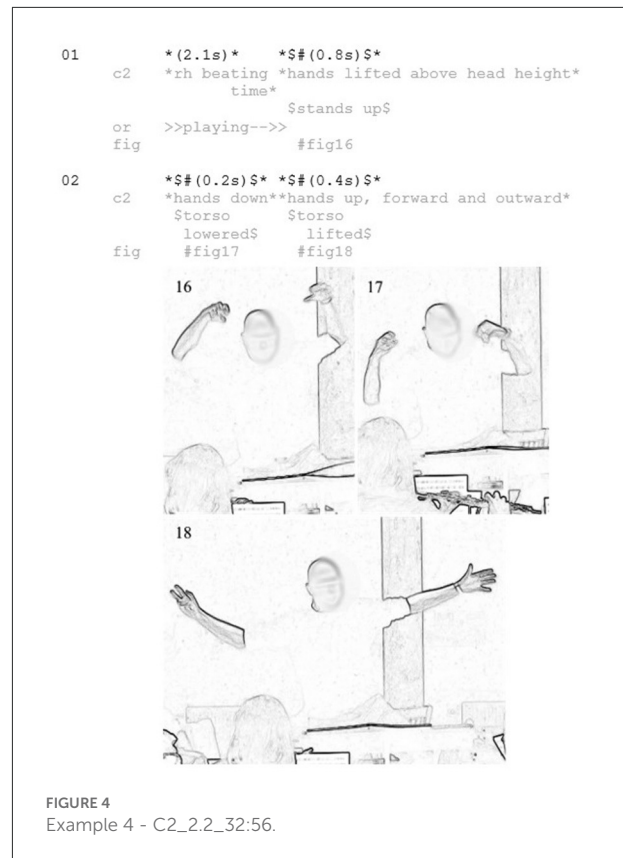


up from his chair, builds up tension by pulling both arms up and toward himself to then lower them slightly and release the tension right at the point of climax with a composite movement upward, forward and outward thus depicting an intense increase in volume.

Figure 5 shows the instruction for a *diminuendo*, using the opposite movement directions along all axes.

Orienting toward an instrument section to his left, the conductor extends his left arm away from his body in preparation with a flat hand facing palm down. He then moves it downward, toward himself sagittally and to the right horizontally, thus representing a coordinated movement along all three axes metaphorically expressing a decrease in volume. At the same time, he continues to beat time with his right hand, which lets us deduce the salience of the movement with the left hand for dynamics as a specific instruction embedded into the broader activity of conducting.

Regarding viewpoint, Figures 4, 5 display similar aspects we already identified in previous fragments. In both examples the



conductor's body is a point of reference for the trajectory of the depicted sound. While in Figure 4, sound is like an OBJECT expanding, almost bursting out and thus moving away in all directions from the conductor, Figure 5 contains the *ex-negativo* meaning construal (see Figure 3), depicting a softer sound as traveling toward the conductor's body.

In this section, we have shown how the much-discussed conceptualization of dynamics (often conceived of as LOUDNESS) as expansion in SIZE can be nuanced with aspects of force and different viewpoint phenomena. While the

mapping of increasing intensity onto the expansion in space away from the conductor's body forms a clear pattern in our data, we now turn to some excerpts, whose gestural imagery does not seem to fit the construal patterns identified thus far.

Complicating the picture: Specificity and viewpoint

Next to the mapping of increasing intensity as movement away from the conductor's body along the vertical, horizontal and sagittal axis, either in isolation or in compound movements, some examples in our data urge for a more nuanced analysis of the metaphorical projections and interactional phenomena at play in orchestra instruction.

For instance, when looking at movement produced along the vertical axis, in certain cases, musical notes which need to be played more strongly and loudly are marked gesturally by a vertically downward movement, which appears to be the exact opposite to the patterns described in the previous section. This, then, raises the question whether these examples are to be categorized as exceptions.

With the following example, we will demonstrate that a more fine-grained analysis of the target domain in terms of more specific (sub)concepts allows to qualify these 'deviant' cases as instantiations of a coherent construal pattern all the same, situated along the very same vertical axis.

In the course of the rehearsal, [Figure 6](#) directly precedes [Figure 3](#) discussed in the previous section. In [Figure 6](#), the conductor suggests an adjustment to a previously performed musical passage, asking the musicians to play a small accent on each beat (line 01), in this case the first of four notes, of the bars in question. This accent is conceptualized visually by pointing the right index finger down onto the extended left index. Following his verbal instruction, the conductor depicts the musical passage by vocalizing the sequence (line 02), audibly stressing each beat while simultaneously repeatedly moving his right index down with each of these vocalized accents. He initiates a playing sequence (line 03) for this fragment and repeats the previous movement on the fly, this time accompanied by audible exhalations with each accent (line 04).

The example raises our interest in several respects. Increasing intensity is conceptualized as a downward instead of an upward movement as observed in [Figure 5](#). To our knowledge, in the above-mentioned studies on musical dynamics in conducting, the depiction of increasing volume with this directionality has not been described as a pattern. We suspect that one reason for this might be that [Poggi \(2017\)](#) and [Opazo \(2018\)](#) explicitly did not include accents as a subcategory of their target domain. [Watson \(2012, p. 170\)](#), however, in a case study on one conductor, identifies a lifted hand that is then dropped to signal a rapid increase in volume.

```

01  C4  en *probeer es een klein accentje*#*per per tel*
    c4  *right index down on *hands held*
    fig  left index x4* #fig22

02  C4  *((vocal.)#BAdeedadee BAdeedadee BAdeedadee DAB)*
    c4  *right index down on left index x4-----*
    fig  #fig23

03  C4  *twee drie EEN*
    c4  two three one
    c4  *both hands beating time*

04  C4  ε*hh° hh° hh° hh°* *(0.5s)ε
    c4  *right index down *hands held-->
    or  on left index x4*
    c4  $playing----->ε

05  C4  nog ene keer*
    c4  one more time
    c4  ----->*

```

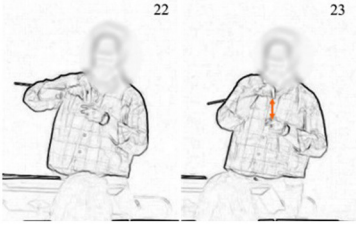


FIGURE 6
Example 6 - C4_1.1_05:41.

Comparing [Figure 6](#) to [Figures 1, 4, 5](#), it appears that the directionality of the movement being depicted along a vertical PATH, is reversed. This construal of accentuation raises the grounding image of a force representing gravity or another downward force, as it is exerted upon an object, as noted by [Boyes Braem and Bräm \(2000, p. 155\)](#). Hence, a musical accent, which typically stands out by a sudden in- and decrease of volume and intensity, tends to be metaphorically construed by a rapid downward movement of the conductor's hand. This construal of downward forces, often in combination with a forward directionality, seems very much at odds with the LOUDNESS IS SIZE metaphor described in most of the examples above, where a gradually increasing sound volume tends to be depicted as an object increasing in size.

In light of the laws of physics and the corresponding logic applied by force metaphors, the depiction of softer, unaccentuated sounds as located spatially higher than louder, accentuated sounds is also visible in [Figure 6](#), when the conductor repeatedly pulls his right index finger up in between the downward accents. In general, when a note is accentuated, this implies the notes after it are to be played at a lower intensity, unless indicated otherwise. So, inherently, by contrast, the note immediately following an accent, will sound softer and less intense, which can be mirrored visually as a bouncing back or retracting motion by the conductor. Immediately following [Figure 6](#), in [Figure 3](#), we observe a retracting movement of that kind as a way to conceptualize softer sounds. When taking another look, we can see that, within seconds, the conductor

changes the movement axes along which he conceptualizes similar aspects of dynamics. While in Figure 6, he uses a straight downward movement to depict an accent, just seconds later, to highlight an aspect of the previous performance, he shifts to the more sagittal movement as shown in Figure 3, away from his body for an accent and toward his body for a *diminuendo*. The construal of force is still present in Figure 3, where the movement toward the conductor's body is performed at shoulder height, while the accent is being depicted by a more rapid movement leading away from the body along a downward gravitational path.

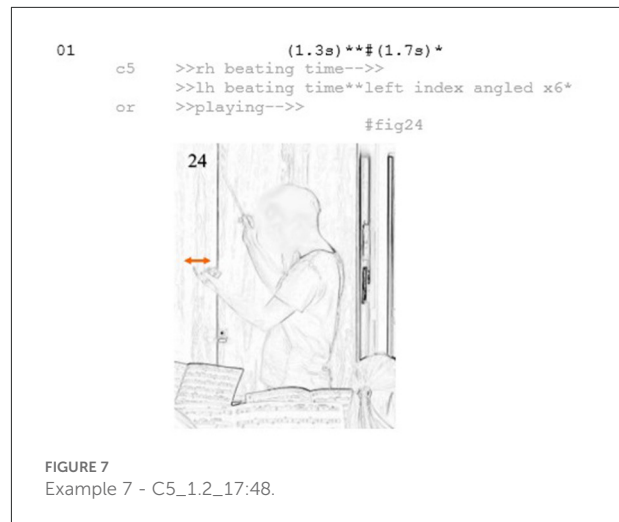
In contrast to Figure 3, Figure 6 also contains a change in viewpoint for the way sound volume is conceptualized in relation to the conductor's body. In Figures 2–5, sound is depicted as if traveling away or expanding more or less away from the conductor. However, in Figure 6, as in Figure 1, we see the notes depicted from an external viewpoint, not relating to the conductor's body as an imagined origin of sound. Therefore, we can see that the conductor's body is not always the point of reference for sound traveling in space. The conductor's gesture space still serves as the point of reference, purely due to the affordances (Gibson, 1979) of the human body. However, sound is depicted as an independent value vis-à-vis the conductor's body as center.

Looking back on Figures 1–6, it seems that while the direction on the vertical axis is interchangeable, a sagittal movement toward the conductor's body in combination with either vertical direction will mostly refer to the performance of a softer sound or a decrease in volume, while a movement away from the conductor's body will mostly refer to a louder sound or an increase in volume. However, there seems to be at least one notable exception to this pattern.

To close off the empirical part of this contribution, we turn our attention to a last excerpt, in which the direction of the sagittal movement seems to run counter to the construal patterns described above, again foregrounding the importance of taking into account viewpoint phenomena.

The example in Figure 7 contains an instruction on the fly, directed at a specific part of the orchestra. In relation to the default conducting position, the conductor's body is rotated to the right, so that he is not facing the camera. At first, the conductor beats time with both hands, but then, as he directs his gaze to a particular subsection of the orchestra, only his right hand continues to beat time. His left hand is directed toward the same subsection, facing palm up with the index finger extended. The conductor then repeatedly pulls his index toward himself, as if asking the musicians to move toward him.

This type of gesture occurs in different variations in our corpus, sometimes involving more than one finger being angled toward the conductor's body or showing an oscillation on the wrist. The latter is also described by Poggi (2017, p. 40) in relation to its meaning in everyday Italian communication,



where it is used to encourage someone to approach the gesturer. In orchestra conducting, this movement is attributed a specific local meaning, namely to increase sound volume, and therefore for musicians to 'come forward' with their sound production in relation to others. The underlying metaphor in this case would be related to physical proximity in the sense that something that approaches you becomes louder (Cox, 2016, p. 98). Although in this example the metaphorical construal of objectified sound traveling along a SOURCE-PATH-GOAL schema remains intact, the depicted directionality of the conceptualized movement is reversed on the sagittal axis. In the gesture in Figure 7, the source of the sound is conceptualized as the musicians producing the sound situated on a path leading toward the conductor who represents the goal.

By producing this wave-like gesture, the conductor exploits his unique aural experience, as the only member of the orchestra who, standing in this central position, can hear the full breadth and depth of the sound (Parton, 2014, p. 408). Therefore, a specific sub-concept of dynamics is put in focus, namely the notion of *balance*, which pertains to the relative sound volume of different individuals and subgroups across the orchestra. This, crucially, implies a viewpoint shift. In Figures 2–5 the aspects of musical dynamics under concern are being construed from the assumed internal perspective of the musicians, using the conductor's body as the reference point for depicting the trajectory of the objectified sound traveling through space. Alternatively, as demonstrated in Figures 1, 6, aspects of dynamics may be construed from an external viewpoint. What makes excerpt 7 stand out against all previous examples is the conductor unambiguously performing a movement from his own viewpoint being physically and prototypically located in front of the orchestra. What this example demonstrates, is the need to actively integrate the construal mechanism of viewpoint in the analysis of this and other examples in order to obtain a

fully motivated account of the visual component structuring this instructional usage event.

The excerpt in Figure 7 very much puts the notion of intersubjective construal on display, featuring the conductor as playing with various viewpoint options during rehearsal, while at the same time—relying on various elements of common ground (Clark, 1996)—assuming the musicians to be able to successfully interpret his various, sometimes fast-changing construal options.

The opposite of this wave-like gesture toward the conductor, which would be a movement away from the conductor or a gesture that stops the imagined approach of sound toward the conductor to indicate a softer or more restrained sound production, did not emerge as a clear pattern in our data. However, another study presents empirical evidence for this variant in the form of a flat hand with the palm facing the musicians and moving toward them instructing them to hold back (Poggi, 2017, p. 42) which suggests that the adoption of this viewpoint also holds for both movement directions on the sagittal axis.

Zooming out, our analysis has demonstrated that a coherent systematic account of a conductor's movements representing aspects of musical dynamics requires taking into account interactional, situationally grounded resources such as the spatial setting and placement of (sub)groups of musicians, as well as (pre-)conceptual and intersubjectively aligned construal mechanisms such as different types of metaphorical mappings and perspectival projections.

Summary

In this contribution we analyzed conductors' movements as they are used to express aspects of musical dynamics, thereby scrutinizing the use of space and spatial relations between conductor and orchestra. Movement direction along three axes was the ultimate analytical focus of this endeavor. We will summarize our findings in terms of the identified movement direction patterns as they can be motivated by underlying construal mechanisms.

With reference to the vertical axis, we can say that the conceptualization of louder sounds as up and softer sounds as down is a clear pattern in our data (Figures 1, 4, 5), confirming earlier studies (Poggi, 2017; Opazo, 2018). However, we also found instances of louder sounds being depicted by conductors in a downward movement just like the representation of softer sounds in an upward movement (Figures 3, 6). We argue that this may relate to different qualitative features of the acoustic experiences under concern, thus rendering them inequivalent target concepts within the overarching domain of musical dynamics. It appears, for example, that instances of a rapid, accent-like increase in volume tend to be depicted more often with a downward movement. On the sagittal axis, most

often, movement away from the conductor translates to louder sounds, whereas movement toward the conductor indicates softer sounds (Figures 3–5). Here as well, our data provide instances in which the opposite directionality, expressed by a wave-like gesture moving toward the conductor, co-occurs with the conductor requesting musicians to play a louder sound (Figure 7). On the horizontal axis, finally, we have found that movement outward from the conductor refers to louder sounds and inward movement to softer sounds (Figures 2, 4, 5). For this axis, more than for the other two, the constant situational re-orientation of the conductor's body as well as the overlap with other aspects that may be conceptualized horizontally (e.g., the sequentiality of the written score) didn't allow us to identify other patterns that would suggest opposite directionality.

Importantly, we see that in no way a one-on-one mapping of a musical dynamic meaning onto a particular axis (horizontal, vertical, sagittal) or even a directionality (up vs. down, away from vs. toward body) on one specific axis could be identified. Depending on the type of experienced loudness, be it a gradual increase or a sudden accent, the movement's directionality, even on the same axis, may alter. Still, the directional co-occurrence patterns we were able to identify allow us to draw conclusions about some of the construal mechanisms underlying these multimodal instructions regarding sound volume and intensity.

The metaphorical mapping LOUDNESS IS SIZE is a dominant pattern, as shown in previous studies. Applying this metaphor, louder sounds are depicted as higher up vertically (Figures 1A,B) and further away from the body both horizontally (Figure 2) and sagittally (Figures 3, 4). Oftentimes, the conductor's body serves as a point of reference from which increasing sound is depicted as traveling further into space, away from the body. Importantly, we observe that the instances of movement toward the conductor's body or center of their gesture space in order to express softer sounds (Figures 3, 5) can only be understood *ex-negativo* on the background of the dominant co-occurrence pattern of outward movement being used to represent increasingly louder sounds.

The structural impact of the schematic SOURCE-PATH-GOAL metaphor is apparent in movements expressed on both the vertical (Figure 1) and sagittal axis (Figure 3), whose directionality may also be reversed (Figure 7). Depending on the viewpoint being adopted, different starting and end points of the traveling sound, either integrating the conductor's and the musicians' bodies or not, may be conceptualized. Our analysis has also revealed that the metaphorical conceptualization of sound as an object facilitates the representation of aspects of musical dynamics as a growing or shrinking movement along two or three spatial axes (Figures 2, 4, 5).

Force metaphors surface in different ways when it comes to musical dynamics. There is the notion of hitting or pushing an imaginary object sagittally away from the conductor's body (Figure 3) to conceptualize louder sounds, as a force created

through physical motion. Gravity as omnipresent force of nature also influences the conceptualization of sounds along the vertical axis, depicting louder sounds as downward falling, softer sounds as upward rising movements (Figure 6).

With regard to the construal mechanism of specificity, the examples above demonstrate that varied metaphorical structures that surface in supposedly opposite movement directions ask for a more fine-grained differentiation of the target concept under scrutiny. We have observed that a metaphorical analysis linking all mappings from a source concept to a static schematic target concept like LOUDNESS cannot account for all the gestural imagery in this subdomain of musical dynamics. Accordingly, we have uncovered allegedly conflicting gestural imagery involving the representation of growing sound volume or intensity by means of either an upward (Figure 4) or a downward vertical movement (Figure 6). Yet, when taking LOUDNESS into account as a dynamically construed target concept, which may be specified on multiple levels of granularity, the allegedly opposing metaphors suddenly make sense vis-à-vis one another. It appears then, for instance, that if we consider the factor of temporality as part of the target structure INCREASING LOUDNESS in order to distinguish between a sudden, accent-like increase from a gradual increase of volume, the sudden downward vertical movement corresponds to the former target specification, whereas the upward vertical movement represents the gradual *crescendo*-type of increase of volume.

As mentioned above, the viewpoint from which instructions on dynamics are conceptualized has a direct impact on the depicted directionality. In general, two kinds of viewpoints can be distinguished in our data when it comes to the depiction of musical dynamics. A first option is taking an external perspective, from which sound is depicted as occurring more or less independently from the participants' bodies—always taking into account the affordances and constraints that the human body imposes on movement and gestures. The second option concerns taking the internal viewpoint of a participant. The application of an external viewpoint is expressed in Figures 1, 6, whereas all other examples depict sound in relation to the conductor's (i.e., the gesturer's) body. Within these participant viewpoints further distinctions can be made, since they can either be ascribed to the conductors themselves or (parts) of the orchestra when it comes to the imagined source of sounds traveling through space. However, often it is not clear whether conductors conceptualize their own body as the imagined source of the sound, therefore mirroring the musicians as the actual source of the sound, or if the conductors merely depict the sound as they themselves envision it, irrespective of it originating from a specific location. It is hard to tell whether it is the conductor's own viewpoint or rather that of the musicians assumed by the conductor in a process of intersubjective coordination. To disambiguate these possibilities, the interactional context or iconic movement for playing a certain instrument can help.

In one type of movement that we described above, conductors unambiguously adopt their own viewpoint in instructions on musical dynamics (Figure 7). With this wave-like gesture, as if inviting the musicians to approach the conductor, which has been described in earlier studies, sound is clearly conceptualized as traveling from the musicians toward the conductor.

Although the different possibilities to perform instructions about musical dynamics are complex and not always clearly distinguishable, their interpretation apparently does not seem to pose any problems to musicians during rehearsal. Rather, there are several factors that contribute to the interpretation of a movement and aid the process of disambiguation. On the one hand, there is the musical score that serves as a reference for participants. On the other hand, additional formal aspects of movement such as handshape and orientation of the hand or qualities such as amplitude or tension feed into the interpretation of movements as part of the usage event. Also, as Poggi (2017, p. 39) notes, aspects of musical dynamics can be expressed either as “global gestures” or by a single parameter within one movement.

Even if, as in our corpus, we cannot exactly monitor or test musicians' reactions to or understanding of these instructions, the observation whether and, if so, how certain passages are repeated or resumed may provide an indication as to whether previous instructions might have been unclear. Additionally, a perception study on instructions about musical dynamics has shown that there is no significant difference in understanding of these movements between participants with and without musical experience (Poggi et al., 2021, p. 1493). An expert group was able to give more fine-grained interpretations (e.g., between a *crescendo* and a *forte*), but general patterns (*louder*, *softer*) were equally recognized by non-experts. This confirms the observation we make, that deeply embodied metaphorical mappings underlie the instructions on musical dynamics, as proven in their local occurrence in our data.

With regard to the method adopted in this study, there are several adjustments that would improve future research. First of all, the segmentation of cases could be conducted from a more economical vantage point, stopping annotation with the first occurrence of a specific movement direction connected to a certain aspect of dynamics across conductors instead of marking all cases of dynamics instructions. Secondly, integrating the musical score could add another layer to the analysis, which we explicitly excluded for this contribution (see Section *Materials and method*). Lastly, inter-coder-agreement tests could enhance the reliability of annotations. However, the goal of this contribution was not an exhaustive account of all occurrences of instructions pertaining to dynamics in our data set, enabling us to generalize our findings. Rather, we aimed at addressing mechanisms underlying movement patterns that surface in our data to enhance our understanding of movement-based communication.

Discussion

Zooming out to a higher level of both methodology and description, finally, this study presents a clear case of the way in which a multimodal analysis of face-to-face interaction—in our case conductors communicating with their orchestra during rehearsal—benefits from a combined analytical approach, in which both cognitive construal mechanisms and situationally bound interactional resources are taken into account (see, among others, Deppermann, 2012; Zima and Brône, 2015). More strongly even, our study has shown that an adequate analysis of the movements made by conductors in depicting aspects of musical dynamics requires an integrated account of both perspectives. Although not the main focus of our current analysis, several of the discussed examples have revealed the relevance of the situational setting of the conductor vis-à-vis the orchestra just as well as the temporal sequence of the actions and usage events as important interactional resources, capable of (co)motivating the kinesemiotic characteristics of the movements expressing aspects of musical dynamics. With regard to the spatial elaboration of a movement, accordingly, it does matter which (sub)sections of the orchestra the conductor is addressing, how deep and broad the rehearsal room is, or which movement a conductor has performed during a preceding interactional sequence, etc.

In this contribution, our main focus has been on the identification of movement direction patterns and the discussion of multiple construal mechanisms underlying and motivating them. By focusing our analysis on the three spatial axes (vertical, horizontal, and sagittal) along with the directionality of the movement on either of them, our empirical analysis has contributed to a better and motivated understanding of the spatial dimensions used to depict elements of musical dynamics. As we scanned our corpus for multimodal (verbal and gestural) instructions pertaining to musical dynamics, we were able to identify three major construal mechanisms (metaphor, specificity, and viewpoint) underlying them. As such, uncovering the power of these construal mechanisms reveals the ways in which locally situated interaction may be embedded in schematic patterns of embodied conceptualization. Yet, in our study we have moved beyond the scope of a metaphorical analysis of multimodal musical expressions anchored to a generalized conceptual metaphor(s) like LOUDNESS IS SIZE, which may be found to subsume the verbal and gestural variations in the corpus. As a matter of fact, our analysis has revealed and overcome two issues, which may render a traditional conceptual metaphor analysis, if taken by itself, rather idle.

The first issue concerns the coarse-grained level of description, at which metaphor analyses sometimes aim to identify a relevant target concept or domain (like LOUDNESS), for which then structural mappings from specific source

domains are being described. As we have shown in our analysis, the frequently quoted, highly schematic concept of LOUDNESS may not be the optimal level of description on which one starts looking for systematic metaphorical mappings in both verbal and gestural expressions. The key to a motivated understanding of the allegedly opposing metaphorical construals serving the same schematic target concept lies in the levels of granularity at which a target concept may be construed and analyzed. In our case, a static and rather schematic representation of the target concept as MUSICAL DYNAMICS or LOUDNESS does not render an accurate and consistent analysis on the part of the source concepts being metaphorically projected. What is required, then, is a dynamic and more fine-grained specification of the target concept, for example, by taking the parameter of suddenness into account when qualifying the experience of an increase in a sound's strength or loudness. Along these lines, the different metaphorical construals of a gradual *crescendo* involving a vertical upward movement, on the one hand, and a suddenly increased volume and intensity in accentuation represented by a downward movement, on the other, can be perfectly motivated. Generally speaking, this means that an accurate analysis of metaphorical construal requires research scrutiny of both the source and target domain thus integrating the construal mechanism of specificity with respect to the target concept in the methodological apparatus.

The second issue pertains to the crucial integration of the analysis of metaphor with the aspect of viewpoint as an intersubjective construal operation thus acknowledging concepts like the theory of mind, intersubjectivity and common ground as core mechanisms in the process of meaning making and, by doing so, bringing the interactional dimension of perspectivization at the core of the construal analysis. In order to arrive at an analysis capable of motivating the directionality of certain movements on a spatial axis, taking into account the viewpoint is a crucial analytical aspect. As we have demonstrated above, a conductor's gestural act to invite a (sub)group of musicians to play louder may result in opposing movements on the sagittal axis, a movement toward the conductor or a movement leading outwards, depending on the adopted viewpoint. Whereas the former originates in the genuine conductor's perspective, the latter, outward movement implies the conductor assuming the musicians' viewpoint, which seems like a reasonable, affordance-related thing to do. Beyond the choice to take either of the interactants' viewpoints, a conductor may also decide to adopt an external perspective, thus representing the entire interaction unfolding in the gesture space in front of them. Crucially, the integration of viewpoint as an omnipresent intersubjective construal operation, enables a motivated and versatile analysis of musical dynamics in terms of its gestural depiction involving different directionalities on various spatial dimensions.

It goes without saying that several conceptual dimensions and interactional resources, like the role of metonymic projections, but also the hand shape, movement of the fingers, etc. have been left out of the current analysis. Yet, as our major point, we hope to have made clear with the excerpts discussed in this paper, the analytical necessity of integrating both interactional and conceptual aspects of the usage event under concern in order to arrive at an even more encompassing multimodal interaction analysis.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: the raw video data used in this study cannot be made available for privacy reasons. Anonymized fragments referred to in the analysis are available in the [Supplementary material](#). Requests to access these datasets should be directed to KM, katharina.meissl@kuleuven.be.

Ethics statement

The studies involving human participants were reviewed and approved by Social and Societal Ethics Committee - KU Leuven. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

KM, PS, and KF conceptualized the study, contributed to the design of the article and the analysis, wrote parts of the theoretical background, and revised the manuscript. KM annotated the data, selected the excerpts from the corpus, wrote the introduction and the first version of the paper, and wrote the methods section. KM and KF wrote the analysis and the discussion. All authors contributed to the article and approved the submitted version.

References

- Antović, M. (2019). "The role of movement in musical signification: from cognitive to conceptual semantics of music," in *Proceedings Volume from the Conference Musica Movet: Affectus, Ludus, Corpus* (University of Belgrade).
- Barlow, M., and Kemmer, S. (2000). *Usage-Based Models of Language*. Stanford, CA: CSLI.
- Boyes Braem, P., and Bräm, T. (2000). "A pilot study of the expressive gestures used by classical orchestra conductors," in *The Signs of Language Revisited: An Anthology to Honor Ursula Bellugi and Edward Klima*, eds K. Emmorey and H. L. Lane (Mahwah, NJ: Lawrence Erlbaum Associate), 143–167.
- Bressem, J., and Müller, C. (2014). "A repertoire of German recurrent gestures with pragmatic functions," in *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK)*, eds C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, and J. Bressem (Berlin; München; Boston, MA: De Gruyter Mouton), 1575–1591. doi: 10.1515/9783110302028
- Bröne, G. (2010). *Bedeutungskonstitution in verbalem Humor: ein kognitiv-linguistischer und diskurssemantischer Ansatz*. Frankfurt am Main; New York: Peter Lang.

Funding

This study reported in the present paper was conducted as part of the MUST: Multimodal Stance Taking in Interaction project (C14/20/040) funded by KU Leuven.

Acknowledgments

We would like to thank the two reviewers who enabled us to improve our article based on their detailed feedback. Also, we kindly thank Fien Andries, Clarissa de Vries, and Valentijn Prové, members of the MIDI research group at KU Leuven, for useful feedback on drafts of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcomm.2022.986733/full#supplementary-material>

SUPPLEMENTARY VIDEO

Example 1a, Example 1b, Example 2, Example 3, Example 4, Example 5, Example 6, Example 7 - respectively for the videos.

- Casad, E. H. (1995). *Cognitive Linguistics in the Redwoods: The Expansion of a New Paradigm in Linguistics*. Berlin; Boston, MA: De Gruyter Mouton.
- Chafe, W. (1994). *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: University of Chicago Press.
- Cienki, A. (2017). "Analysing metaphor in gesture: a set of metaphor identification guidelines for gesture (MIG-G)," in *The Routledge handbook of metaphor and language* *Routledge Handbooks in Linguistics*, eds E. Semino and Z. Demjén (London; New York, NY: Routledge), 131–147.
- Cienki, A. (2022). The study of gesture in cognitive linguistics: how it could inform and inspire other research in cognitive science. *WIREs Cogn. Sci.* 2022, e1623. doi: 10.1002/wcs.1623
- Clark, H. H. (1996). *Using Language*. West Nyack: Cambridge University Press.
- Clark, H. H. (2016). Depicting as a method of communication. *Psychol. Rev.* 123, 324–347. doi: 10.1037/rev0000026
- Cox, A. (2016). *Music and Embodied Cognition: Listening, Moving, Feeling, and Thinking*. Bloomington: Indiana University Press.
- Croft, W., and Cruse, D. A. (2004). *Cognitive Linguistics*. Cambridge: Cambridge University Press.
- Dancygier, B. (2017). *The Cambridge Handbook of Cognitive Linguistics*. Cambridge: Cambridge University Press.
- Deppermann, A. (2012). How does 'cognition' matter to the analysis of talk-in-interaction? *Lang. Sci.* 34, 746–767. doi: 10.1016/j.langsci.2012.04.013
- Dolscheid, S., Shayan, S., Majid, A., and Casasanto, D. (2013). The thickness of musical pitch: psychophysical evidence for linguistic relativity. *Psychol. Sci.* 24, 613–621. doi: 10.1177/0956797612457374
- Eitan, Z. (2013). "How pitch and loudness shape musical space and motion," in *The Psychology of Music in Multimedia*, eds S.-L. Tan, A. J. Cohen, S. D. Lipscomb, and R. A. Kendall (New York, NY: Oxford University Press), 165–191.
- Feldman, J., Epstein, D., and Richards, W. (1992). Force dynamics of tempo change in music. *Music Perception* 10, 185–203. doi: 10.2307/40285606
- Feyaerts, K. (2013). "A cognitive grammar of creativity," in *Creativity and the Agile Mind: A Multidisciplinary Approach to a Multifaceted Phenomenon*, Vol. 21, eds T. Veale, K. Feyaerts, and C. Forceville (Berlin: Mouton de Gruyter), 205–227. doi: 10.1515/9783110295290.205
- Feyaerts, K., and Oben, B. (2014). "Tracing down schadenfreude in spontaneous interaction. evidence from corpus linguistics," in *Schadenfreude: Understanding Pleasure at the Misfortune of Others*, eds W. W. van Dijk and J. W. Ouwkerk (Cambridge: Cambridge University Press), 275–291.
- Feyaerts, K., Oben, B., Lackner, H. K., and Papousek, I. (2017). Alignment and empathy as viewpoint phenomena: the case of amplifiers and comical hypotheticals. *Cogn. Linguist.* 28, 485–509. doi: 10.1515/cog-2016-0109
- Gehrrens, K. W. (2006). *Music Notation and Terminology*. Project Gutenberg. Available online at: https://www.gutenberg.org/files/19499/19499-h/19499-h.htm#CHAPTER_XIII (accessed April 11, 2022).
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Givón, T. (2005). *Context as Other Minds: The Pragmatics of Sociality, Cognition and Communication*. Amsterdam; Philadelphia, PA: John Benjamins Publishing Company.
- Goodwin, C., and Salomon, R. (2019). Not being bound by what you can see now. Charles Goodwin in conversation with René Salomon. *Forum: Qualitative Social Research*, 20, 11.
- Haddington, P., Keisanen, T., Mondada, L., and Neville, M. (2014). *Multiactivity in Social Interaction: Beyond Multitasking*. Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Hsu, H.-C., Brone, G., and Feyaerts, K. (2021). When gesture 'takes over': Speech-embedded nonverbal depictions in multimodal interaction. *Front. Psychol.* 11, 1–23. doi: 10.3389/fpsyg.2020.552533
- Hutchby, I., and Wooffitt, R. (2008). *Conversation Analysis*. 2nd ed. Cambridge; Malden, MA: Polity.
- Ivaldi, A., Sanderson, A., Hall, G., and Forrester, M. (2021). Learning to perform: a conversation analytic systematic review of learning and teaching practices in performing arts lesson interactions. *Learn. Cult. Soc. Interact.* 28, 1–17. doi: 10.1016/j.lcsi.2020.100459
- Janzen, T. (2022). Embodied cognition: ASL signers' and english speakers' use of viewpoint space. *Lang. Cont.* 22, 227–258. doi: 10.1075/lic.00020.jan
- Johnson, M. (1987). *The Body in the Mind: The Bodily Basis of Meaning, Imagination and Reason*. Chicago, IL: University of Chicago Press.
- Johnson, M. (2017). *Embodied Mind, Meaning, and Reason: How Our Bodies Give Rise to Understanding*. Chicago, IL: University of Chicago Press.
- Johnson, M. L., and Larson, S. (2003). "Something in the way she moves"-metaphors of musical motion. *Metaphor Symbol* 18, 63–84. doi: 10.1207/S15327868MS1802_1
- Kok, K. I., and Cienki, A. (2016). Cognitive grammar and gesture: points of convergence, advances and challenges. *Cogn. Linguistics* 27, 67–100. doi: 10.1515/cog-2015-0087
- Kövecses, Z. (2015). *Where Metaphors Come From: Reconsidering Context in Metaphor*. New York, NY: Oxford University Press.
- Ladewig, S. H. (2020). *Integrating Gestures: The Dimension of Multimodality in Cognitive Grammar*. Berlin; Boston, MA: De Gruyter Mouton.
- Ladewig, S. H., and Bressem, J. (2013). New insights into the medium hand: discovering recurrent structures in gestures. *Semiotica* 2013, 203–231. doi: 10.1515/sem-2013-0088
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago, IL: University of Chicago Press.
- Lakoff, G., and Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.
- Langacker, R. W. (1987). *Foundations of Cognitive Grammar: Theoretical Prerequisites*. Stanford, CA: Stanford University Press.
- Langacker, R. W. (1991). *Foundations of Cognitive Grammar: Descriptive Application*. Vol. 2. Stanford, CA: Stanford University Press.
- Langacker, R. W. (2008). *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.
- Langacker, R. W. (2010). *Grammar and Conceptualization*. Berlin; New York, NY: De Gruyter Mouton.
- Larson, S. (2012). *Musical Forces: Motion, Metaphor, and Meaning in Music*. Bloomington: Indiana University Press.
- Maiorani, A. (2020). *Kinesemiotics: Modelling How Choreographed Movement Means in Space*. New York, NY: Routledge.
- Meissl, K., Sambre, P., and Feyaerts, K. (Submitted). "Contrast pairs in orchestra instruction as a window into the multimodal expression of stance," in *Papers of the Linguistic Society of Belgium* 16.
- Messner, M. (2020). Gesangliche Demonstrationen als instruktive Praktik in der Orchesterprobe. *Gesprächsforschung Online-Zeitschrift zur verbalen Interaktion* 21, 309–345. Available online at: <http://www.gespraechsforschung-online.de/fileadmin/dateien/heft2020/instruktionen-theater.pdf>
- Mittelberg, I. (2017). Embodied frames and scenes: body-based metonymy and pragmatic inferencing in gesture. *Gesture* 16, 203–244. doi: 10.1075/gest.16.2.03mit
- Mittelberg, I. (2018). Gestures as image schemas and force gestalten: a dynamic systems approach augmented with motion-capture data analyses. *Cogn. Semiotics* 11, 1–21. doi: 10.1515/cogsem-2018-0002
- Mondada, L. (2019a). Contemporary issues in conversation analysis: embodiment and materiality, multimodality and multisensoriality in social interaction. *J. Pragmatics* 145, 47–62. doi: 10.1016/j.pragma.2019.01.016
- Mondada, L. (2019b). *Conventions for Multimodal Transcription*. Available online at: <https://www.lorenzamondada.net/multimodal-transcription>
- Opazo, P. (2018). *The Moving and Expressive Body: A Study on the Semiotic Resources Used by Classical Orchestra Conductors and Their Metaphorical Associations in the Presentation of Musical Dynamics* (master thesis). Vrije Universiteit Amsterdam, Amsterdam.
- Parton, K. (2014). Epistemic stance in orchestral interaction. *Soc. Semiot.* 24, 402–419. doi: 10.1080/10350330.2014.929389
- Parton, K., and Edwards, G. (2009). *Features of Conductor Gesture: Towards a Framework for Analysis Within Interaction*. Sydney, NSW.
- Poggi, I. (2002). "The lexicon of the conductor's face," in *Language, Vision and Music. Selected Papers from the 8th International Workshop on the Cognitive Science of Natural Language Processing, Galway, Ireland 1999*, eds P. McKeivitt, S. Ó Nualláin, and C. Mulvihill (Amsterdam, Philadelphia: John Benjamins Publishing Company), 271–284.
- Poggi, I. (2017). Signals of intensification and attenuation in orchestra and choir conduction. *Normas* 7, 33. doi: 10.7203/Normas.7.10423
- Poggi, I., and Ansani, A. (2016). Forte, piano, crescendo, diminuendo. Gestures of intensity in orchestra and choir conduction. in *Proceedings of the 4th European and 7th Nordic Symposium on Multimodal Communication (MMSYM 2016)*, eds P. Paggio and C. Navarretta (Copenhagen), 111–119. Available online at: <http://www.ep.liu.se/ecp/contents.asp?issue=141> (accessed June 4, 2022).

- Poggi, I., D'Errico, F., and Ansani, A. (2021). The conductor's intensity gestures. *Psychol. Music* 49, 1478–1497. doi: 10.1177/0305735620963179
- Prové, V., and Feyaerts, K. (2022). *Pitch Metaphors and the Body in Singing Classes*. CogniTextes 22. doi: 10.4000/cognitextes.2037
- Rohrer, P. L., Vilà-Giménez, I., Florit-Pons, J., Esteve-Gibert, N., Ren, A., Shattuck-Hufnagel, S., et al. (2020). *The MultiModal MultiDimensional (M3D) Labelling Scheme for the Annotation of Audiovisual Corpora*. Stockholm. Available online at: <http://marcs.uws.edu.au/links/ICoMusic09/index.html>
- Sambre, P. (2021). Taking the trumpet up there: enactment of embodied high pitch in a multimodal body schema. *Linguistics Vanguard* 7. doi: 10.1515/lingvan-2020-0114
- Sambre, P., and Feyaerts, K. (2017). Embodied musical meaning-making and multimodal viewpoints in a trumpet master class. *J. Pragmat.* 12, 10–23. doi: 10.1016/j.pragma.2017.09.004
- Schuldt-Jensen, M. (2015). What is conducting? Signs, principles, and problems. *Signata* 383–421. doi: 10.4000/signata.1126
- Shayan, S., Ozturk, O., and Sicoli, M. A. (2011). The thickness of pitch: crossmodal metaphors in Farsi, Turkish, and Zapotec. *Senses Soc.* 6, 96–105. doi: 10.2752/174589311X12893982233911
- Simon, S., and Feyaerts, K. (2020). *Conducting Fanfare Orchestras. A Multimodal Corpus*. KU Leuven MIDI.
- Sousa, G. D. (1988). *Musical conducting emblems: An investigation of the use of specific conducting gestures by instrumental conductors and their interpretation by instrumental performers* (doctoral thesis). Ohio State University, Columbus, Ohio. Available online at: https://etd.ohiolink.edu/apexprod/rws_etd/send_file/send?accession=osu1217257892&disposition=inline (accessed June 4, 2022).
- Spitzer, M. (2018). Conceptual blending and musical emotion. *Musicae Scientiae* 22, 24–37. doi: 10.1177/1029864917714302
- Stoeckl, H., and Messner, M. (2021). Tam pam pam pam and mi – fa – sol: constituting musical instructions through multimodal interaction in orchestra rehearsals. *Multimodal Commun.* 10, 193–209. doi: 10.1515/mc-2021-0003
- Sweetser, E. (2012). “Introduction: viewpoint and perspective in language and gesture, from the Ground down,” in *Viewpoint in Language: A Multimodal Perspective*, eds B. Dancygier and E. Sweetser (Cambridge: Cambridge University Press), 1–22.
- Sweetser, E., and Sizemore, M. (2008). “Personal and interpersonal gesture spaces: Functional contrasts in language and gesture,” in *Language in the Context of Use*, eds A. Tyler, K. Yiyoung, and M. Takada (Berlin, New York: De Gruyter), 25–51.
- Szczepek Reed, B., Reed, D., and Haddon, E. (2013). NOW or NOT NOW: coordinating restarts in the Pursuit of Learnables in vocal master classes. *Null* 46, 22–46. doi: 10.1080/08351813.2013.753714
- Talmy, L. (1988). Force dynamics in language and cognition. *Cogn. Sci.* 12, 49–100. doi: 10.1207/s15516709cog1201_2
- Talmy, L. (2003). *Toward a Cognitive Semantics*. Cambridge, MA: MIT Press.
- Taylor, J. R. (2002). *Cognitive Grammar*. Oxford: Oxford University Press.
- Thiemel, M. (2001). *Dynamics*. Oxford Music Online - The Oxford Dictionary of Music. Available online at: <http://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000008458> (accessed October 13, 2022).
- Tolins, J. (2013). Assessment and direction through nonlexical vocalizations in music instruction. *Null* 46, 47–64. doi: 10.1080/08351813.2013.753721
- Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.
- Traugott, E. C., and Dasher, R. B. (2001). *Regularity in Semantic Change*. 1st Edn. Cambridge: Cambridge University Press.
- Verhagen, A. (2015). “Grammar and cooperative communication,” in *Handbook of Cognitive Linguistics HSK Handbuecher zur Sprach- und Kommunikationswissenschaft/Handbooks of Linguistics and Communication Science (HSK)*, eds E. Dabrowska and D. Divjak (Berlin; Boston, MA: De Gruyter Mouton), 232–252.
- Veronesi, D. (2014). Correction sequences and semiotic resources in ensemble music workshops: the case of Conduction[®]. *Null* 24, 468–494. doi: 10.1080/10350330.2014.929393
- Veronesi, D., and Pasquandrea, S. (2014). Doing (things with) sounds: introduction to the special issue. *Soc. Semiot.* 24, 369–380. doi: 10.1080/10350330.2014.929379
- Watson, C. (2012). *Gesture as Communication: The Art of Carlos Kleiber* (doctoral thesis). Sydney Conservatorium of Music - The University of Sydney, Sydney. Available online at: <https://ses.library.usyd.edu.au/handle/2123/8797>
- Watson, C. (2019). *The Craft of Conducting – A General Introduction, Vol. 4. Music & Practice*. Available online at: <https://www.musicandpractice.org/volume-4/the-craft-of-conducting-a-general-introduction/>
- Weeks, P. (1985). Error-correction techniques and sequences in instructional settings: toward a comparative framework. *Hum. Stud.* 8, 195–233. doi: 10.1007/BF00142993
- Weeks, P. (1996). A rehearsal of a beethoven passage: an analysis of correction talk. *Res. Lang. Soc. Interact.* 29, 247–290. doi: 10.1207/s15327973rlsi2903_3
- Whiten, A. (1991). *Natural Theories of Mind: Evolution, Development, and Simulation of Everyday Mindreading*. Oxford: B. Blackwell.
- Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., and Sloetjes, H. (2006). “ELAN: a professional framework for multimodality research,” in *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*, 1556–1559.
- Zbikowski, L. M. (2002). *Conceptualizing Music: Cognitive Structure, Theory, and Analysis*. New York, NY: Oxford University Press.
- Zbikowski, L. M. (2017). *Foundations of Musical Grammar*. New York: Oxford University Press.
- Zima, E., and Brône, G. (2015). Cognitive linguistics and interactional discourse: time to enter into dialogue. *Lang. Cogn.* 7, 485–498. doi: 10.1017/langcog.2015.19