



# Rhythm and the embodied aesthetics of infant-caregiver dialogue: insights from phenomenology

Kasper Levin<sup>1</sup>  · Maya Gratier<sup>2</sup>

Accepted: 27 February 2024  
© The Author(s) 2024

## Abstract

This paper explores how phenomenological notions of rhythm might accommodate a richer description of preverbal infant-caregiver dialogue. Developmental psychologists have theorized a crucial link between rhythm and intercorporeality in the emergence of intersubjectivity and self. Drawing on the descriptions of rhythm in the phenomenology of Maurice Merleau-Ponty, Erwin Straus, Henri Maldiney and Maxine Sheets-Johnstone, the paper emphasizes the role of art and aesthetic processes proposing that they not only be considered as metaphorical or representational aspects of rhythm but as primary resources that can enrich and deepen our understanding of self-emergence and intercorporeality in preverbal infant-caregiver dialogue.

**Keywords** Infant-caregiver · Intersubjectivity · Rhythm · Aesthetics · Vitality affects

## 1 Introduction

At first sight the preverbal child's movements can appear uncoordinated and incidental. However, as developmental research has suggested over the last decades, the preverbal child's movements express embodied intentions which some consider to be fundamental or constitutional to the development of consciousness, communication and intersubjectivity. As Pérez and Español (2016) have pointed out in a review of multimodality in infant research, the study of movement in infant-caretaker interaction was considered a core developmental aspect of the pioneering work of schol-

---

✉ Kasper Levin  
kaslevin@ruc.dk

<sup>1</sup> Department of People and Technology, Roskilde University, Roskilde, Denmark

<sup>2</sup> Laboratoire Ethologie Cognition Développement (LECD), Université Paris Nanterre, Nanterre, France

ars like Condon and Sander (1974), Trevarthen et al. (1977) Tronick (1978), Daniel Stern, Beatrice Beebe and Joseph Jaffe (Beebe et al., 1979; Jaffe et al., 1973), but has not since been followed up systematically. As Pérez and Español argue, the multimodal dimension of movement was abandoned and exchanged for an increased focus on the linguistic aspects of interaction. In this context, they point out that the pioneering studies of movement in the 1970s were primarily aimed at two closely interrelated aspects: the temporal structure or organization of exchanges, and the identification of multimodal units based on the segmentation of movement. While the first aspect refers predominantly to a quantitative approach aimed at measuring a relationship in terms of the sequential distribution of events, the second is mostly directed towards identifying the qualitative emotional meaning or significance emanating from movement-based expression.

What Pérez and Español only mention in passing is that the early pioneers' qualitative work on describing movement in interaction is permeated by a language borrowing heavily from the arts. In a recent chapter on the development of what they term a multimodal perspective of infant-caregiver interaction, Español et al. (2022) more directly express how "This stream of thought opened the doors to an aesthetic view of development that paid attention to a temporal interaction dimension whose flow and meaning are similar to the musical phenomenon" (2022, p. 6). As they point out, the notion of music as an aesthetic process in relation to preverbal expression and communication is implicitly or explicitly a prominent characteristic within a lot of developmental research on preverbal infancy.

However, making this implicit quality explicit raises several epistemological and ontological questions pertaining to the role of aesthetic qualities in early human development. On the one hand, it is worth thinking about how an aesthetic sense might actually be involved in the formation and transformations of the earliest social relationships, and on the other hand it is worth thinking about how our knowledge of the arts and aesthetic processes in particular might inspire new angles for approaching the study of early social interaction.

In the observational studies of infant-caregiver interaction that flourished in the 1970s, recurring and pervasive musical concepts, such as harmony, synchrony, attunement and rhythm, could be considered as "opening the doors" (Español et al., 2022, p. 6) to qualitative aspects of infant-caregiver interaction. In particular, rhythm was described as a crucial multi-, cross- or amodal aspect of communication (see Davis, 1982). In that context, the musical notion of rhythm offered a means to understand the extent to which bodily movement could convey meaning. A major inspiration behind much of the work focused on rhythm in social interaction put forth by scholars such as Beebe, Stern and Jaffe (Beebe et al., 1979; Jaffe et al., 1973), Trevarthen (1977) and others, was the paper by Lashley (1951) on hierarchically organized cognitive motor plans preceding voluntary movement in behavior. Condon and Sander's groundbreaking study (1974) showing that neonates' movements are synchronized with caretaker's speech provided the basis for thinking of rhythm as a distributed intermodal phenomenon. While Lashley's studies of rhythm were informed by a behavioral perspective, Condon and Sander's study emphasized the linguistic aspects of the relationship between bodily coordination and sound. Bateson (1975) and Trevarthen (1977) described the vocal coordination between adults and young infants as

protoconversation, also emphasizing the linguistic resemblances between nonverbal and verbal communication.

What was attractive about the concept of rhythm to Stern and his colleagues was not simply linguistic or behavioral, but rather the promise of an approach that both captured the qualitative role of a pre-reflective, and therefore non-representational, temporal dynamics and provided an opportunity for the segmentation of behavior. Both Condon (1982) and Stern (1977) described face-to-face social interaction, between adults or between adults and infants, as a coordinated ballet where each partner's movements are synchronized and attuned to the other's.

Stephen Malloch and Colwyn Trevarthen later recast earlier work on interaction and protoconversation as a dance-like or musical form of communication,

When the infancy researchers reported discoveries of delicate expressions and sensitive responses passing between young infants and their mothers, they described it in terms of rhythmic patterns of engagement that could be represented as 'musical' or 'dance-like'. (2009, p. 1)

The main idea here is that the time-based artforms of music and dance become means of capturing basic experiential qualities of movement across different modalities in time and form, which are notoriously difficult to capture in conventional conceptualizations of perceptual experience. Trying to approach this elusive domain of perceptual experience, scholars such as Beebe, Jaffe, Feldstein, Stern and others conducted a range of studies concerning the qualitative dynamics of rhythmic coordination, synchrony and emotional attunement in caretaker-infant interaction (Beebe et al., 1979, 2000; Jaffe et al., 1973, 2001). In these accounts, rhythm in interaction is defined as a unifying recurrent temporal pattern that organizes expression across different modalities. For example, this is evident in the empirical finding that "mother-infant vocal rhythms are highly correlated with those of looking, head movement, and gesture" (Jaffe et al., 2001).

These multimodal explorations were followed up by scholars like Trevarthen, Malloch, Gratier and others (Delafield-Butt & Trevarthen, 2015; Delavenne et al., 2008; Devouche & Gratier, 2001; Gratier, 2009; Gratier & Apter-Danon, 2009; Gratier & Trevarthen, 2007; Jover & Gratier, 2023; Malloch, 2000; Malloch & Trevarthen, 2009; Trevarthen, 2009). Although these more recent empirical studies do not primarily focus on the multimodal aspects of movement, as highlighted by Pérez and Español (2016), they do indeed affirm that infants' early communication should be understood as amodal structures and rhythmic dynamics akin to music. In essence, these studies have provided convincing support as well as a theoretical framework for considering rhythm as a primary source and motive for intersubjective communication. A psycho-biological perspective is at the core of the theory of 'communicative musicality', which assigns a prominent role to rhythm. Accordingly, music in its broadest sense is seen as a fundamental organizer of human communication and cognition through evolution (Trevarthen, 2000). Musicality is thus an inherent feature of human neuro-cognitive, perceptual and emotional capacities. In this sense, it is not a learned ability but rather an innate potential for embodied relationality that gives rise to a capacity for prospectively controlled movement and a rhythmic social agency.

In this paper we propose that a phenomenology of aesthetics can further enrich a description and understanding of the role of rhythm in embodied movement, consciousness and intersubjectivity. The aim of this paper is to specifically address the relationships between the phenomenological aesthetics of rhythm and early embodied social engagement by exploring key insights in the works of Daniel Stern, Maurice Merleau-Ponty, Erwin Strauss, Henri Maldiney and Maxine Sheets-Johnstone.

## 2 Daniel Stern and metaphors of affect from dance and music

When thinking about the experience of young infants, one frequently falls back on metaphors of music. This is hardly strange, given that both types of experience are nonverbal, are intuitively grasped, and involve the flow of forms in time. I will suggest that certain basic experiences of time and form that are common to our encounter with music are also common to an infant's ordinary, daily, socio-affective interactions. (Stern, 2000, p. 21)

Stern's as a phenomenological theory of developmental psychology, takes us further into the work of philosophies of perception and movement and how they work might shed new light on current research taking us forward into understanding the complex dynamics of human relating from the earliest stages of cognitive development.

In his earliest work, Stern highlighted both the aspects associated with the temporal dynamics of movement in interaction and aesthetic models involving concepts from music and dance. As Stern mentions in his book *The First Relationship* (Stern, 1977), some of his earliest collaborators were dancers and choreographers, who shared an interest in the microanalysis of movement. Stern ultimately sees the infant-caregiver interaction as a "dance choreographed by nature" (1977, p. 3). Thus, his focus on the expressive qualities of infant-caregiver interaction can be seen as stemming from a prior interest in the temporal dynamics of dance. As Stern himself asserts: "After so much observation of the micro-local level of mother-infant interaction, metaphors from music and dance not only crept into my writing but also became a way for me to think about what I saw" (Stern, 1977, p. 23). To Stern, one of the strengths provided by musical phenomena for the description of interactional processes was that they captured the ongoing qualities of shared affect beyond the view that interaction depends on a static matching body schema through behavioral imitation. "For there to be an intersubjective exchange about affect," Stern argues, "then, strict imitation alone won't do" (Stern, 1985, p. 139). What the musical metaphor brings to the table in this context is a relational view of interaction and affect that highlights the amodal and 'between-ness' dimensions of interaction. The fundamental psychological idea that Stern (1977) puts forth is that "Behaviors, thoughts, feelings, actions have a musical quality. Each behavioral, or affective, or even cognitive phrase—that is, the shortest meaningful chunking—has a contour in time" (1977, p. 213).

In this sense it could be argued that Stern's own characterization of the musical concepts as metaphors is misleading. The musical conceptualization of movement contours, which Stern comes to call "vitality affects" or "activation contours" (Stern,

1985), does not mean that contours are per se musical or “like” music, but rather that *vitality affects* conceptualize a phenomenon which had until then only been described in music or dance as auditory or visual phenomena. While qualities such as happy, sad, angry, surprised etc. appear as clear linguistically sharable categories, ‘vitality affects’ are experienced and understood at a pre-reflective amodal level, which cannot be shared as a subjective or linguistic category. Vitality affects indeed stem from feeling shared in intersubjective engagement, as distributed experiences extending over two or more bodies. Psychologically, Stern operationalizes this level as ‘the emergent self’, linked to special qualities expressed in dynamically changing affective states, motivation, attunement and emotional tensions. For Stern’s understanding of early infant development and communication, it is significant that these affective qualities are available to infants, but remain immanent to all human relations unfolding in categorial affects, language and communication. As he argues:

If a certain gesture by the mother is to be “correspondent” with a certain kind of vocal exclamation by the infant, the two expressions must share some common currency that permits them to be transferred from one modality or form to another. That common currency consists of amodal properties. There are some qualities or properties that are held in common by most or all of the modalities of perception. These include intensity, shape, time, motion, and number. Such qualities of perception can be abstracted by any sensory mode from the invariant properties of the stimulus world and then translated into other modalities of perception (Stern, 1985, p. 152).

As Stern points out here, our primary communication with the world and other beings must emerge from a common amodal and pre-reflectively shared foundation that allows for a translation from one modality to another. Pre-reflective experience refers to the embodied, immediate and non-conceptual way in which we perceive and engage with the world prior to any reflective thoughts. Amodal does not mean cross-modal, which is the transfer from one modality to another. Rather, Stern describes amodal as the direct relational forces of events in their intensity of happening, not peculiar to any perceptual modality, but common to all of them. However, linguistically, our common vocabulary for affective states or emotional categories does not fit the amodal intensities of vitality affects, as Stern describes, “These elusive qualities are better captured by dynamic, kinetic terms, such as ‘surging,’ ‘fading away,’ ‘fleeting,’ ‘explosive,’ ‘crescendo,’ ‘decrescendo,’ ‘bursting,’ ‘drawn out,’ and so on” (1985, p. 54). Rather than emotional categories or feeling states, these processual percepts are better apprehended through an aesthetic language borrowed from the arts. When Stern writes in *The Interpersonal World of the Infant*, “Abstract dance and music are examples par excellence of the expressiveness of vitality affects” (1985, p. 56) he believes that a musical composition that moves you affectively is a good example of the unmediated relations of intense forces constituting the processual formation of emotions. However, it could be argued that the musical concepts go beyond the representative function of metaphors. One might consider the aesthetics of music as a processual rethinking of the psychological concept of affect. On this account, Stern’s musical conceptualizations are not more metaphoric than vitality contours

are. Music might be an immediate or direct singular disposition of the human body in expressive movement, as Malloch and Trevarthen (2009) have posited. From this perspective, the musical aspects of infant-caregiver interaction appear as an actualization of an expressive relational reality linking movement to affect. Rhythm here serves to conceptualize the amodal unification of the expressive forces in movement.

For instance, a rhythm, such as “long short” (– –), can be delivered in or abstracted from sight, audition, smell, touch, or taste. For this to occur, the rhythm must at some point exist in the mind in a form that is not inextricably bound to one particular way of perceiving it but is rather sufficiently abstract to be transportable across modalities. It is the existence of these abstract representations of amodal properties that permits us to experience a perceptually unified world. (Stern, 1985, p. 152)

As a temporal phenomenon saturated with vitality contours that are transportable across modalities, rhythm is thought to work as a “common currency,” both of the earliest emergence of a unified self and of the primary intersubjectivity of affective attunement and communication with a caregiver. However, the ontological status of rhythm as an aesthetic image remains vaguely described as something presupposed that must “exist in the mind” and “in a form” without any explicit explanation of how rhythm can both be a constituting force of the emerging self and already existing *in* the mind. This presupposition of rhythm or other aesthetic phenomena as something simply given in the emergence of self and intersubjectivity returns us to the challenge associated with the prevailing belief, not only in Stern’s work, but also in the pioneering research tradition he participated in, that musical concepts and even art and aesthetic qualities serve as metaphors for affect.

Much developmental research on early interaction has focused on the communicative aspects of rhythm. In this context the project of relating rhythm to expression can be traced back to a much older hypothesis according to which language and music are part of the same system and share the same evolutionary or biological origin<sup>1</sup>. This hypothesis has primarily been pursued through a linguistically informed transmission model of communication based on systematic observations and experiments that focus on quantifying musical aspects of vocalization in infant-caregiver dialogue. As Jaffe and colleagues describe, “speech rhythm is one easily quantified index of the rich communicative package that mothers and infants display and coordinate in face-to-face communication” (Jaffe & Feldstein, 1970, p. 8). In other words, the easily quantifiable vocalizations in face-to-face transactions are seen as being part of a far more fundamental system of dialogic temporality. Research within this tradition has led to interesting and valuable findings as well as new insights for child development. However, the scientific effort to quantify interactional processes, often reduces the very musical concepts which “opened the door” to a rich study of the dynamics of affective emergence and dialogue.

<sup>1</sup> As Rousseau argues in his *Essay on the Origin of Languages*: “cadence and sounds are born together with syllables: passion rouses all of the [vocal] organs to speech and adorns the voice with their full brilliance; thus verse, song, speech have a common origin” (Rousseau, 1997, p. 286 [1781])

Thus, besides forgoing the project of capturing the multimodal aspects of movement in interaction, research from the 1980s onwards has largely replaced qualitative conceptualizations with mechanistic statistical approaches. For example, the central role of rhythm in perinatal development has been operationalized and successfully confirmed through temporal measures of sensory-motor coordination (Provasi et al., 2014). Recent studies of infant-caregiver dialogue largely focus on delimiting cross-modal relations between, for example, bodily movement, touch, gaze and vocalization (Bourjade et al., 2023; Cordes et al., 2017; Egmoose et al., 2018; Elmlinger et al., 2023) rather than on the amodal dynamics of movement. In other words, studies on infant-caregiver interactions have been intent on describing and quantifying binary or pre-defined structures such as those involving turn-taking relations (Gratier et al., 2015) which deflect from the question of how movement impulses interpersonal coordination and which qualities of movement are involved in attaining synchrony and coordination. Thus, cohort-based studies that seek to measure aspects of social interaction, though they do provide valuable partial insights, lose sight of the dynamic and wholistic dimension of expressive movement unfolding in time.

In this sense, developmental science has not benefitted from a deep understanding of the significance of rhythmic and dynamic phenomena for human relating. These concepts have remained metaphoric and secondary to the understanding of self and intersubjectivity. Thus, the difficulty of bringing the qualitative descriptions of *dynamic forces* of intersubjective expression into the task of operational and quantifiable axiomatic categories for modelling communicative process has never really been overcome. Some attempts have been made to identify improvisational processes between musicians based on insights from infant sociality (Gratier et al., 2017; Gratier & Magnier, 2012). Stern was very much aware of the problems associated with transferring the artistic processes of shaping and refining expressivity of performances into technical language. As he describes,

... what distinguishes an artistic performance from a technical one. The difference is one of elastic rhythms versus formal rhythms. The magic of elastic rhythms is in the precise shaping of the vitality affects to express the exact feelings that are behind the acts transmitted to the audience. And the magic in a therapy session or in intimate relations, beyond or underneath the explicit meanings, also lies there. This is where authenticity resides. (Stern, 2004, p. 57)

The distinction between elastic and formal rhythms however continues to pose a problem for formalizing, coding, notating or categorizing rhythm. Yet, the idea of an affective “authenticity” behind an artistic performance may hold some promise for empirical enquiry. The resort to “magic” in Stern’s description points to a linguistically liminal association between a relational ethos and a relational aesthetics. Interestingly, Trevarthen (2004) theorized the continuum between morality and beauty as played out in everyday spontaneous parent-infant engagement.

In his last book *Forms of Vitality* (2010) Stern takes up the intricate relationship between the relational development of affect and the qualitative aspects inherent in the arts in a more direct manner, arguing that the arts “serve as a good example of what happens in daily interactions between two people” (2010, p. 76). Based on this

observation he goes on to describe in further detail how different systems for coding qualitative dimensions of artistic performances have been developed within the fields of music, dance and theater which are all time-based arts. Although Stern sometimes expresses rather naïve or romantic naturalizations of art, for example in the characterization of the caregiver as an artist (Stern, 1977, p. 159), his main point remains relevant — beyond the conventional and linguistically-fixed structure of signs, there appears to be an aesthetic domain fundamental to the qualitative understanding of intersubjective relationality that needs to be studied.

Stern explicitly appeals to another possibility for pursuing research based on insights from the arts in both *The Present Moment* (2004) and *Forms of Vitality* (2010). With reference to the philosophy of Edmund Husserl and of Maurice Merleau-Ponty, he instigates a new phenomenological psychology of human relational development based on the study of affect in movement. Stern asserts that phenomenology is generally relevant because “It provides a starting place to look for vitality forms of the feel of being alive” (2010, p. 34). Indeed, vitality affects are seen as experiential phenomena that do not fall within the categories of perception or sensation per se. He states that “To advance, we may need a multiple approach combining neuroscience, phenomenology, and some ideas from the world of the arts with its particular perspective on dynamic forms and vitality dynamics” (2010, p. 27). Stern’s own intuition of advancing the concept of vitality affects through a phenomenological approach combined with neuroscience and the arts is interesting because it calls for a more thorough reexamination of the qualitative aspects of embodied perception and expressive movement based on a cross-disciplinary approach that assigns a phenomenological and epistemological value to art.

### 3 Merleau-Ponty and the phenomenological embodiment of rhythm

The embodied phenomenology of Maurice Merleau-Ponty seems to share a fundamental affinity with studies focused on rhythmic aspects of movement in developmental psychology. Insisting on the fundamentally aesthetic nature or “style” of the primary “motor intentionality” immanent to our embodied capacity for movement, Merleau-Ponty wrote that “The body cannot be compared to the physical object, but rather to the work of art. In a painting or in a piece of music, the idea cannot be communicated other than through the arrangement of color or sounds” (2012, p. 152). This aesthetic conception of embodiment, challenges the traditional way of thinking about both intersubjectivity and communication. Opposing the conventional idea of expressive movements as abstract representations of inner processes of reflection or thought, Merleau-Ponty’s suggestion is that there is a primordial expressive aspect immanent to the very formation of subjective experience. Merleau-Ponty believes that “[all] perception, all action which presupposes it, and in short every human use of the body is already primordial expression” (2007, p. 267). The perception of the world is always already shaped or “styled” through invisible or imperceptible movements immanent to sensation. Consequently, the expressive act of self-movement becomes the limit of communication. In other words, because self-movement or gesture mark an instance of a radical unified and pre-reflective consciousness of move-



ment, “The consciousness of my gesture, if it is genuinely an undivided state of consciousness, is no longer the consciousness of a movement at all, but rather an ineffable quality that cannot teach us about movement” (Merleau-Ponty, 2012, p. 543 n. 60). Consequently, the project of capturing this ineffable aspect of movement led him to connect our lived experience to a range of aesthetic or artistic modes of being. As he writes in *Phenomenology of Perception* on the relationship between embodied experience and rhythm: “Insofar as I inhabit a “physical world,” where consistent “stimuli” and typical situations are discovered [...] my life is made up of rhythms that do not have their reason in what I have chosen to be, but rather have their condition in the banal milieu that surrounds me.” (2012, p. 86). In this account, rhythm is described as a pre-reflective and relational phenomenon that “makes up” or binds articulations of ‘stimuli’, representational choices and “typical situations” into the acquired familiarity of ordinary or unremarkable life. As opposed to the commonly held concept of rhythm as an experiential sequence or succession of articulated points of stimuli, rhythm is here presented as a pre-reflective relational field or affective background which I immediately recognize as mine. Thus, in musical terms, rhythm as a qualitative force is rather to be understood as the interval or the imperceptible relation that can be sensed but not heard. It is in the effort of describing this silent or imperceptible backdrop of perceptual experience as such that Merleau-Ponty resorts to aesthetic resources drawn from his engagements with Mallarmé, Valéry, Proust, El Greco, Cézanne and others (see Wiskus, 2015). In this sense, Merleau-Ponty’s descriptions of rhythm holds the same challenges as Stern’s vitality affects in terms of not remaining within a representational understanding of perceptual forms or emotions. As opposed to representation, the logic of rhythm and vitality affects is that of emergence, which means that the expressive quality of a rhythm or vitality affect cannot refer to the existence of an already established form. It is on this point that Merleau-Ponty’s phenomenological description of rhythm is at odds with Stern’s suggestion that “the rhythm must at some point exist in the mind in a form” (Stern, 1985, p. 152). The phenomenological problem is here that perceiving a subjective actualization of a specific rhythm or vitality affect in movement presupposes that our embodied unity and relationality emerges or is expressed from where the rhythm or affect germinates. It is from this perspective that Merleau-Ponty suggests that the body shares a basic existential condition with art,

A novel, a poem, a painting, and a piece of music are individuals, that is, beings in which the expression cannot be distinguished from the expressed, whose sense is only accessible through direct contact, and who send forth their signification without ever leaving their temporal and spatial place. (2012, p. 153)

In other words, like a work of art, the body “is a knot of living significations” (2012, p. 153) that is unified and in relational communication with the world through expressive encounters and arrangements. It is also in this sense that Merleau-Ponty proposes that art and poetry are more complete expositions of the emergent or expressive phenomena immanent to our subjective sense-making in language. In his essay *Le langage indirect et les voix du silence* (1960) Merleau-Ponty more specifically describes how poetic language works as a pre-articulated background for our everyday speech

and representational language. Rhythm in poetic language can generate new affective styles and new meaning by making expressive the silent relations between sign and signification, as opposed to the conventional understanding of language as a direct expression of our inner thoughts. Thus, rhythm in poetry can be seen as a silent voice of embodied or pre-reflective affective processes immanent to representational language. From this perspective, aesthetics become relevant as a means of exploring fundamental processes involved in pre-reflective perception and intersubjectivity.

Merleau-Ponty's concept of rhythm can be connected to enactive approaches to social cognition and participatory sense-making (De Jaegher et al., 2017; De Jaegher & Di Paolo, 2007; Krueger, 2014a, c). Referring to Merleau-Ponty and studies on early infant-caregiver exchanges, the phenomenologist Joel Krueger defends what he calls the "joint ownership thesis" (Krueger, 2013). This suggests that some early affective experiences are structurally composed to be jointly owned by more than one subject. For example, experiences of positive affect in rhythmic musicality in the infant-caregiver relationship can be described as a phenomenological structure constituting an inherently dyadic experience. In this context, Krueger has advocated for the enactivist notion of the "musically extended mind" (Krueger, 2010, 2014b), which is associated with both bodily and environmentally extended emotions (Colombetti & Krueger, 2015; Krueger, 2014d). Krueger's suggestion is closely associated with other contemporary enactive approaches to the phenomenology of music-making and improvisation (Høffding & Schiavio, 2021; Iyer, 2004, 2016; Schiavio et al., 2016; Schiavio & Hoffding, 2015) that combines inspiration from embodied cognition (Gallagher, 2005; Varela et al., 1991) and ecological psychology (Gibson, 1979). While music in this perspective is considered an environmental scaffolding of the mind giving primacy to joint affective experiences, the aesthetic aspect is primarily and explicitly described as a secondary function that articulates, shapes, or regulates natural perception and emotions. In other words, music often remains within a representational image of art and aesthetic composition, presupposing that perception and affective states are naturally given to be shaped or regulated<sup>2</sup>. While this conceptualization effectively explains how infants and caregivers sometimes "do things" with rhythm and *use* music as an "aesthetic technology" (Krueger, 2014b) of affective functions in jointly owned experiences, it does not address the ontological question associated with Stern's suggestion of rhythm as an active immanent aesthetic process in the emergent co-constitution of self and intersubjectivity.

<sup>2</sup> Unlike Merleau-Ponty's explicit primacy of art and aesthetic processes in the constitution of subjective experience, enactive approaches to music and embodiment most often focus on naturalizing the role of musical aesthetics in relation to perceptual and affective givens. As the enactivist phenomenologist Alva Noë argues, "what is true of the experience of the work of art is true of human experience quite generally" (Noë, 2012, p. 2). In his understanding, art or aesthetic compositions presuppose perception as a natural activity given in subjective experience. Although this does make sense from the point of view of a socio-culturally specific subjective experience, it does not describe the active and developmental role of art and aesthetics in the constitution of subjectivity. What is often missing in these phenomenologically embodied and enactive notions of art is an account of the creative and constitutive role of aesthetics in the emergence of subjectivity (Levin, 2016).

#### 4 Erwin Straus and the pathic moment of sensing

The aesthetics of rhythm and the emergent aspects of pre-reflective perception it reveals may further be understood through the distinction made by the phenomenological psychologist Erwin Straus between the *pathic* moment and the *gnostic* moment of perception. As Straus explains, “by the pathic moment, we mean the immediate communication we have with things on the basis of their changing mode of sensory givenness” (1966, p. 12). In contrast to this, the *gnostic* moment describes the reflective recognition or representation in perception. As Straus states “The gnostic moment merely develops the *what* of the given in its object character, the pathic the *how* of its being as given” (1966, p. 12). This distinction is also significant in light of Straus’s primary emphasis on the concept of *sensing* [*Empfinden*]. Inspired by Martin Heidegger’s concept of *Mit-Sein*, Straus describes how perception is a form of knowing which is always preceded by sensing as an experience of *being-with*. What is vital to the difference between perception and sensing is that, as Straus remarks, “sensing has in itself the character of change, and thus a definite temporal structure” (1963, p. 18). It is this temporal quality of perpetual change and movement in sensing that characterizes the pathic moment and its relationship to rhythm.

Long before Daniel Stern’s suggestion of psychological notions of rhythm, vitality affects and attunement, Erwin Straus highlighted the pathic moment as an unmediated intermodal rhythmic unity of phenomenal experience using an aesthetic model in relation to children’s development,

Just as sight, hearing, touch, and taste are interrelated so is sensing as such bound in an *inner* connection to vital, living movement. The music and the movements of a march, the music and the movements of a dance are intermodally united. There are no particular kinds of association which tie motion to sound and rhythm, motion quite immediately follows music. Long before the youngster is taught conventional dance steps, he dances in rings, hops to the hopping movement of a polka, is drawn by the music of a march into the ranks of the marching columns (1963, p. 233).

According to Straus, the infant’s relation to different music styles, or in Stern’s terms, to their vitality affects, reveals an unmediated bond between hearing a rhythm and performing rhythmic movement. The obvious force of rhythm, and perhaps art more generally, as described in the above example, is that its qualitative strength does not belong to a specific perceptual or sensory modality. In this sense, it attracts the effort of describing the amodal or pathic qualities of movement – i.e. actualizing movement as a virtual dimension of the body. This is not restricted to a question of understanding the primacy of rhythm in dance experiences, but rather the primacy of an amodal unity in the connection between movement and sensing. As Straus further explains, “The unity of sensing and moving becomes obvious in the phenomenon of the dance, but it is a unity not limited to this particular case. It encompasses all sensing and all animated movement” (1963, p. 233). Straus compellingly proposes that the aesthetics immanent to the experience of rhythm are constitutive of all self-world relations. As Straus suggests “The sensing subject does not have sensations, but, rather, in his

sensing he has first himself. In sensory experience, there unfolds both the becoming of the subject and the happenings of the world.” (Straus, 1963, p. 351). On this account, rhythm and aesthetic forms in art are not metaphors for sensory processes, but rather organizing forces of self-world relations played out at the sensory level. Can art be directly constitutive, not only of our understanding of our own situated reality, but also of our embodied proprioceptive sensory reality?

## 5 Henri Maldiney and rhythm as the truth of the sensible

A more radical formulation of the immanent relationship between art, rhythm and sensation is found in the work of the French phenomenologist Henri Maldiney who states that, “rhythm is the truth of this primary communication with the world, which essentially consists of αἴσθησις [aísthēsis] from which aesthetics takes its name, the sensation in which feeling is articulated with movement” (Maldiney, 2012b, p. 208). He considers rhythm as the foundation of all art because the role of art is to set in motion affective and perceptual forms to create autonomous worlds. Thus, as he maintains, “Art is the truth of the sensible because rhythm is the truth of ἡ αἴσθησις [aísthēsis]” (Maldiney, 2012b, p. 208). Just as music creates coherent worlds through the tension of different intervals and rhythms in sound, visual art can be said to create autonomous visual worlds through rhythmic variations of tensions in line, color and density. Consequently, rhythm is not only aesthetically central to the perception of sound in music. As Maldiney asserts,

There is no other aesthetics than rhythm. There is no rhythm other than aesthetic [rhythm] [...] Saying that all rhythm is aesthetic is to say that the experience of rhythm – in what we encounter *there* and *how* it “takes place” – is the order of sensing (and the communication in sensing). (Maldiney, 2012b, p. 208).

At the core of Maldiney’s argument is the idea that rhythm, as an experience immanent to all sensation, informs the constitutive process of perceptual experience. Thus, in line with Straus’ description of the pathic moment as “the immediate communication we have with things” (1966, p. 12), Maldiney suggests that rhythm is what gives form to the process of the immediate unfolding of our sensing, without itself being a specific form of perception or sensation. In other words, if rhythm is that by which perception is formed, rhythm cannot itself be identified as a form of perception. As Maldiney argues, “The fundamental error, repeated constantly and everywhere, is the confusion of rhythm with cadence” (Maldiney, 2012a, p. 17). Once a rhythm passes over into the form of a cadence it loses the openness of a rhythmic process of formation. More specifically, according to Maldiney, rhythm concerns the self-organizing process in which the living and the lived become one as a temporal quality (Levin et al., 2019). Going beyond the classical Greek definition of rhythm as “The name for order in movement” (Plato, 1980, p. 46 [665a]), Maldiney insists that rhythm is not simply the order of movement in time, because, as he maintains, “It is not enough that the articulating moments constitute an order, this order also has to carry a temporal dimension” (Maldiney, 2012b, p. 215). Thus, according to Maldiney the phenomenon

of rhythm cannot be subordinated to a form or sequence of extended points, but must be understood as a more processual temporal quality of formation or what he terms “implied time” (2012b, p. 216). Thus, when we confuse rhythm with a cadence or any other rhythmic shape, “The words that express it (interval, regularity, mark, periodicity) all return to metric space and time as spatialized time accentuating the tick-tock of a clock” (Maldiney, 2012a, p. 17). Maldiney wants to draw our attention to rhythm as a primary relational force that cannot be subordinated to its limit points such as the tick-tock of a clock which conventionally conceptualizes rhythm as a shape of articulated points or positions in time. On this account, the tick-tock interval of a clock is not constituted by pre-articulated marks of the tick and the tock. Rather, the tick-tock endpoints emerge or become expressive as limits through the movement immanent to the tension between them. In this sense, the relational quality of sensing is immanent to the temporal quality implied in the movement. As he radically asserts, “A rhythm does not unfold in time and space. It is the generator of its space-time” (Maldiney, 2012a, p. 20).

Maldiney’s philosophical concept of rhythm as a primordially implied relational process appears radical, but it also makes sense in relation to Stern’s psychological description of vitality affects in intercorporeal attunement. Before a dyadic line of communication between two moving bodies can be unfolded or extended in time, there needs to be a relational attunement, i.e. the creation of a common space that allows the bodies to communicate. Stern’s work and that of other researchers suggest that infants are born with a temporal capacity for synchronizing, imitating and supplementing expressive movements, which empirically supports Maldiney’s phenomenological idea of ‘implied time’. In developmental research this implied temporality is predominantly understood as an inner sense of time or what Trevarthen has conceptualized as an “Intrinsic Motive Pulse” (Trevarthen, 2000, 2016). In his theory, however, the concept of primary inner temporality still remains subordinated to the succession of discrete points in the form of a pulse, be it implicit and therefore not directly perceptible. According to Maldiney’s concept of implied time, the qualitative attunement of a common space cannot be understood in terms of its successive points as for example an inner pulse, transactions or turn-taking of the connected subjects’ experiences. Rather, the affective tensions in sensing intercorporeal movements constitute a pre-reflective and amodal field which is put into mutual communication by the generative rhythm born out of different experiential forms. In this sense, implied time means that time is immanent to temporality, not to a subject or an object. Maldiney’s contribution shares the phenomenological point of departure in embodied movement with both Merleau-Ponty and Straus, but his emphasis on rhythm as an autonomous generative force of implied temporality makes his account unique. While both Merleau-Ponty and Straus describe rhythm as a phenomenon mediating between the subjective and objective aspects of embodied experience, Maldiney considers rhythm a more primordial aspect of human existence. As he states, “Rhythm is an existential” (Maldiney, 2012a, p. 22). The existential primacy of rhythm in sensing is connected to the assertion that art is “the truth of the sensible”. It is also in this sense that art can become a promising avenue for renewed dialogues within developmental science.

Although these different philosophical descriptions of rhythm, sensing, movement and aesthetics are useful for grasping the abstract dynamics of affective relations, they also pose an empirical challenge for the study of infant-caregiver communication. If rhythms are to be considered a pre-reflective field of amodal and implied temporality, and this tension of differential forces cannot be subordinated to a given shape or form that is better grasped through art or aesthetic practices, how do we approach them phenomenologically?

## 6 The primacy of movement and dynamic line

One approach to the empirical question of the phenomenological relation between embodied perception in infants, aesthetics and rhythm can be found in the work of the philosopher and dancer Maxine Sheets-Johnstone. In her book *Primacy of Movement* (2011) she asserts that “thinking in movement is an infant’s original mode of thinking” (2011, p. xxv), and in this sense, “movement is our mother tongue” (2011, p. 195). Sheets-Johnstone’s suggestion of a fundamental epistemological primacy of movement is predominantly born out of a critical reassessment of the general phenomenological descriptions of consciousness in the works of Edmund Husserl and Maurice Merleau-Ponty. However, Sheets-Johnstone’s primacy of movement also hinges on a fundamental developmental psychological point of departure. As she argues:

We learn our bodies by moving and in moving both create and constitute our movement as a spatio-temporal dynamic. If we look more deeply into the matter, we discover that movement is the originating ground of our sense-makings, in phenomenological terms, the originating ground of transcendental subjectivity; we constitute space and time originally in our kinesthetic consciousness of movement. (2011, p. 139).

In accordance with the descriptions above from Merleau-Ponty, Straus and Maldiney, Sheets-Johnstone affirms that an understanding of our fundamental perception of space and time, as well as our sense-making processes and intercorporeal communication, must begin in an account of movement. As Sheets-Johnstone’s dictum asserts, “*movement forms the I that moves before the I that moves forms movement*” (2011, p. 119). It is in this context that she has criticized Stern’s conceptualizations of attunement for not making explicit the dynamic concordance between movement and affect because he does not conceive movement as such, but rather as observable action or behavior (2008, p. 201). According to Sheets-Johnstone, our access to movement is double in the sense that we have both perceptions of movement, as an overt *kinetic* aspect, and feelings of self-movement, as a *kinesthetic* feature. Unlike the observations of the kinetic aspects of the motion of objects in an environment or the movement of one’s own body, the kinesthetic aspect of self-movement has a qualitatively different mode of appearance.

In originary self-movement, what is created and what is constituted are one and the same. A further way of putting this fundamental character of self-movement is to say that self-movement is originarily not only *not* an object in the usual sense — a thing that appears; it is by the same token not a phenomenon that endures across different perceptions of it or that has different profiles to begin with. (Sheets-Johnstone, 2011, p. 132)

While the kinetic qualities of movement can be approached through other modalities such as visual, tactile or auditory perception, the kinesthetic qualities of self-movement are moreover transient and only available through movement itself. As Sheets-Johnstone points out, “In self-movement, a particular unfolding dynamic is kinesthetically present that cannot be otherwise kinesthetically present except by our moving differently and thereby creating a different qualitative dynamic” (2011, p. 132). This account obviously has empirical consequences for the qualitative study of kinesthetic dynamics. Since the quality of self-movement does not provide an appearance in front of us to observe, listen to or touch, we are only left with a “reverberating felt sense” (2011, p. 132), which is challenging to capture.

Although Stern’s descriptions of affect attunement through vitality affects or activation contours coincides, to a large extent, with a phenomenological description of these kinesthetic aspects of intercorporeal movement, his empirical focus on the overt and kinetic, according to Sheets-Johnstone, leaves out the descriptions of the phenomenological relation between the kinesthetic and the affective. Movement in Stern’s experimental studies is necessarily reduced to a fixed set of criteria that refer to overt kinetic inter-actions between mother and infant. Sheets-Johnstone further states that:

Attunement might, in effect, continue beyond the experimental protocol, generating a more complex pairing. An infant might, for example, spontaneously respond to its mother not only by repeating its movements but by increasing their intensity, by slowing them down or speeding them up, by making them bigger or smaller, by making movements similar to its mother’s, by transposing the same kinetic dynamics to a totally new modality, and so on (2008, p. 209).

She points out that the spontaneous intercorporeal sense-making in Stern’s experimental studies of affect attunement are limited to, or by, the required definite “stopping-points,” which makes it possible to determine matching criteria that segment the movements in terms of their beginning, middle and end – i.e. acts or behavior. For example, Stern axiomatically defines the criterion that infants must sense the mother’s response to his or her affective expression as a point for jumping into the stream of interaction. As Sheets-Johnstone critically suggests, “there is no reason to think any “stream of interaction” would actually end coincident with the protocol” (2008, p. 209). Affect attunement might not be limited to the infant’s automatic affective expression and the mother’s response, which she considers an adultist perspective that primarily leads to descriptions aimed at explaining human behavior rather than phenomenologically describing and understanding the full meaning of them. For example, the infant might not only sense the mother’s response, but also

respond to it in a way that either elaborates the dynamic of the exchange, or creates a complementary dynamic of attunement, which invites the mother to respond. Thus, in Sheets-Johnstone's phenomenological perspective on movement, what characterizes attunement is not simply a question of a segmented sequence of interaction in an axiomatic bodily expression and response. It is also a certain expressive quality developing in the way the individuals are "*moved to move*" (Sheets-Johnstone, 2008, p. 208) in an open-ended creation of time and space. As she continues: "Through just such spontaneous ongoing harmonious pairings, an infant itself can open the door to further possibilities of attunement, and in the process, to the possibility of developing its capacity for more and more complex pairings" (2008, p. 210). On this background she argues that, although the existing empirical research on joint attention and turn-taking document intersubjective relations in early infancy, they do not address the question of how intersubjectivity is constituted and how intersubjectivity both shapes and is shaped by development (Sheets-Johnstone, 2008, p. 281). To address this, what Sheets-Johnstone suggests is that our understanding of intercorporeal movement should be substantiated by a more phenomenologically open approach that starts "from the bodily logos or affect-logic-kinetic of infancy rather than from assumption-laden adultist perspectives" (2008, p. 211). In this sense, the infant's movements are not only receptive but also creative. From this perspective, naturally embodied dispositions such as turn-taking, joint attention and imitation cannot be considered autonomous functions, but must be explored from their source in "the livingly present, kinetic tactile-kinesthetic bodies" (Sheets-Johnstone, 2000). Accordingly, as a mutually emergent and open-ended phenomenon, the expressive quality of intercorporeal attunement should be studied in a way that takes into account kinetic-kinesthetic dynamics, where sense-making and movement go hand in hand. Or rather, in a situation where "*what is created and what is constituted are one and the same*" (2011, p. 132). What Sheets-Johnstone proposes is that the exploration and understanding of intercorporeal movement in turn-taking should be done with or through the descriptive phenomenology of self-movement rather than pre-established and axiomatic categories of movement observed and described from an outsider's perspective. It can be argued that this qualitative ambition of exploring the kinetic-kinesthetic aspects of turn-taking, joint attention and imitation, as Sheets-Johnstone suggests, is already integral to some of the existent developmental studies of infant-caregiver studies. For example, as already stated above, the qualitative appeal to the creative aspects of musical phenomena such as rhythm in Stern's, Trevarthen and other developmental scholars' conceptualizations of movement, seems to be an opening toward the creative and phenomenological depth that Sheets-Johnstone calls for. And indeed some developmental psychologists have forcefully argued in favor of a second-person approach to understanding intersubjective dynamics, that is through the embodied experience of participants rather than the abstracted descriptions of observers (Reddy, 2008). Nonetheless, the phenomenologically embodied description of music and rhythm has by and large remained either axiomatically formal or rather vague in developmental science.



## 7 Rhythm and the virtual forces of dance

The vagueness, or lack of a clear general definition of rhythm, is not surprising, but rather the natural consequence of the phenomenological primacy of movement, according to Sheets-Johnstone. Although she does not comment on the concept of rhythm in developmental psychology, in her book *Phenomenology of Dance* (1966) she argues that since rhythm does not exist before moving (in her case dancing) then: “To advance a definition of rhythm prior to a determination of what dance is, is to look at dance as already possessing the characteristics to be noted [...] Let us rather begin by looking again at the phenomenon in question, at dance, the creation of virtual force” (1966, p. 200). Rhythm in this sense begins with a phenomenological description of what it means to dance or to move oneself. Sheets-Johnstone’s description of rhythm here comes close to Maldiney’s radically processual concept of rhythm which loses its rhythmic openness and immediate emergent qualities if it is described in terms of a structure or form. And like Maldiney suggests, Sheets-Johnstone finds the truth of this openness in art, or rather in the aesthetic practice of dance. However, unlike Maldiney’s primacy of rhythm in art, Sheets-Johnstone refers to a phenomenological primacy of being “*moved to move*” drawing on the philosopher Susanne Langer’s (1953) concept of art as the creation of illusion. This is not to be understood as illusion in the sense of make-believe, delusion or a deception of reality, but rather the creation of sheer appearance of the sensory qualities removed from their values of practical function. As Sheets-Johnstone affirms,

phenomenologically, dance does create an illusion of force in the appearance of moving forces which are seen not as actual concrete body movements of everyday life, but as virtual forces emanating from the dancer’s body. In fact, the moment concentration is focused upon the actual physical body, and not the movement, the dance disappears. (Sheets-Johnstone, 1966, p. 78)

It is these “virtual forces” that are primary and central to Sheets-Johnstone’s description of rhythm in movement. Rhythm is here described as a “dynamic line” which is “a unique qualitative organization of forces from beginning to end” (1966, p. 88). As a rhythm, a dynamic line is elusive in the sense that it only exists in the creation or performance of the dance. Although the rhythmic structure of a dance can indeed be notated and repeated through a series of points or intervals in a movement, the dance does not exist outside the felt quality of the dynamic flow of virtual forces inherent to it. With reference to Straus, the dynamic line of virtual forces could be described as a pathic moment of perception. What is intuited in the rhythm of a dance is not the reflective knowledge of the structure in the counting of meter or mastering the series of connected points and proper accents, but rather the qualitatively felt actualization of the virtual forces of the dynamic flow of the movement. As Sheets-Johnstone describes, “the dance does not come alive until the dancer passes beyond a mastery of the structure, and comes to realize the dynamic flow inherent in the total piece” (1966, p. 109). Sheets-Johnstone’s analysis of dance makes a compelling case for an aesthetic primacy in qualitative research. In *The Primacy of Movement*, she observes that,

The quality of a work is not simply the formal aspect of a technique, an aspect we can notice and analyze reflectively in terms of the way in which an artist has kinetically engaged and utilized materials. Quality is coincident with the created form itself; it is there, present in the work, and is immediately apparent in genuinely aesthetic experiences of the work. (Sheets-Johnstone, 2011, p. 102).

This is how Sheets-Johnstone's phenomenological contribution through the aesthetic practice of dance can help develop a more qualitatively rich conception of rhythm. Thus, in relation to the references to rhythm in developmental psychology and Stern's approach to infant-caregiver interaction as a "dance choreographed by nature" (1977, p. 3), Maxine Sheets-Johnstone specifies why this aesthetic orientation is not coincidental. The description of the forces involved in the qualitative organization of movement in dance can bring us closer to a richer understanding of the rhythm inherent to movement in infant-caregiver dialogue.

## 8 Cardinal structures of rhythm

As a dancer, Sheets-Johnstone explores the experiential structures or qualities of virtual forces in movement. Through what she considers to be a Husserlian introspective method, consisting of the performance of epochée and eidetic variation, she goes on to reduce the phenomenon of movement to four qualitative primary cardinal structures (tensional, linear, amplitudinal and projectional) constituting the overall qualities of energy, space and time.

These qualitative aspects — dynamic structures inherent in movement — enter into and define our global qualitative sense of any particular movement variation; they make all of the variations immediately distinctive to us as variations. (Sheets-Johnstone, 2011, pp. 122–123)

These qualities make up the fundamental invariant or eidetic structures of kinesthetic consciousness. In her phenomenological approach, she describes the *tensional* quality that refers to the sense of effort or intensity, which we experience clearly, for example, when we lift, pull or push something. We typically experience tension as a feeling of tautness, hardness or limpness. *Linear* quality describes the total configuration of the moving force. Thus, the linear quality is felt as the contour or directional path of the body's movements when, for example, something is sensed as curved, direct, twisted or vertical. The *amplitudinal* quality is experienced through the spatial extensiveness or constrictedness (sic.) of our body, as when we feel expansive in reaching for something or contracted when ducking. The *projectional* category covers the explosive, abrupt, sustained, ballistic or other qualitative modes in which we feel the release of force or energy. One of the strengths of Sheets-Johnstone's phenomenology of movement is that the categories are immediately associated with time and space. Consequently, the linear and the amplitudinal describe spatial aspects of movement, while the tensional and the projectional describe temporal aspects of movement. Conversely, one could say that the experience of space in movement is

produced from the combination of linear and amplitudinal qualities and the experience of time in movement is produced through tensional and projectional qualities (Sheets-Johnstone, 2011, p. 123). The fact that the four different qualities are inherent in our general experience of space and time also means that they are not reducible to empirically distinct movements, but are always inherently connected in the global experience of movement. The four qualities are thus only analytically revealed in their complex experiential relation to the other qualities. As Sheets-Johnstone describes:

the linear quality of any movement does not exist apart from the tension required to project the line, the area displaced in creating the line, nor the manner in which the line is projected. All qualities of movement are internally bound to one another in and through movement, in and through force which is the global structure of the presented phenomenon (Sheets-Johnstone, 1980, p. 51).

Consequently, specific movements cannot be pointed out as being, for example, purely linear, projectional or tensional. Rather, all four qualities are specific elements in the complex dynamic totality of movement. It is in this sense that rhythm or the dynamic line is fundamental to the understanding of the specific qualitative structuring of movement. Rhythm here is a felt unfolding dynamic of these cardinal structures of embodied consciousness. As already mentioned, Sheets-Johnstone points out (2008) that her phenomenologically informed cardinal structures are, to a large extent, compatible with Stern's illustrative description of the virtual forces in affect attunement. However, Sheets-Johnstone's phenomenological insistence on the unique quality of self-movement rather than action or behavior opens a potential path for moving further towards fleshing out the full richness and potential of intercorporeal rhythms in infant-caregiver dialogue.

## 9 Conclusion

Based on our readings of phenomenological theories in philosophy and psychology, we have argued in this paper that rhythm should be conceived as an aesthetic dimension of infant-caregiver dialogue. We suggest that insights from the arts and from the subjective and intersubjective aesthetic processes they afford, can enrich and frame the understanding of the constituents of the developing self and its concomitant intersubjective relations. The phenomenological aesthetics of rhythm, in various fields of human expression, artistic and social, may prove useful in fleshing out the rather mysterious temporal processes that underlie human social connectedness in general, and in particular at the very start of life. Understanding the complex phenomenon of rhythm as it appears in the emergence of intercorporeal dialogue in a developmental context should be recognized as a major field of enquiry in the cognitive sciences today and as one that can realistically be addressed empirically as well as conceptually. There are rich phenomenological resources for initiating this conceptual exploration in dialogue with current enactive approaches to social cognition, participatory sense-making, and music perception. However, unlike classic explorations of the phenomenological aesthetics of rhythm by Merleau-Ponty, Maldiney, and

Straus, contemporary studies often ignore or downplay the constitutive role of art and aesthetics for subjectivity and the development of sociality. Empirically, this could be addressed not only through more holistically oriented studies in developmental psychology, in interdisciplinary dialogue with musicians, dancers, and artists but also by conducting comparative research across developmental psychology and the arts.

**Acknowledgements** We thank Bjarne Sode Funch and Denise Newman whose comments helped to improve the paper.

**Author contributions** Both authors made substantial contributions to the conception or design of the work, drafted or revised it. Both authors approved the submitted manuscript and agree to be accountable for the accuracy and integrity of the content.

**Funding** Open access funding provided by Roskilde University

**Data availability** Not applicable.

## Declarations

**Ethical approval** Not applicable.

**Informed consent** Not applicable.

**Research involving human participants and/or animals** Not applicable.

**Competing interests** None of the authors have competing interests to disclose in relation to the submitted manuscript.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bateson, M. C. (1975). Mother-infant exchanges: The epigenesis of conversational interaction. *Annals of the New York Academy of Sciences*, 263(100), 101–113. <https://doi.org/10.1111/j.1749-6632.1975.tb41575.x>.
- Beebe, B., Stern, D. N., & Jaffe, J. (1979). The kinesic rhythm of Mother-Infant Interaction. In A. W. Siegman, & S. Feldstein (Eds.), *Of speech and time: Temporal Speech patterns in interpersonal contexts* (pp. 23–34). Lawrence Erlbaum Associates.
- Beebe, B., Lachmann, F., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2000). Systems models in Development and Psychoanalysis: The case of vocal rhythm coordination and attachment. *Infant Mental Health Journal*, 21(1–2), 99–122.

- Bourjade, M., Dafreville, M., Scola, C., & Jover, M. (2023). Six-month-old infants' communication in a comparative perspective: Do maternal attention and interaction matter? *Journal of Experimental Child Psychology*, 231, 105651. <https://doi.org/10.1016/j.jecp.2023.105651>.
- Colombetti, G., & Krueger, J. W. (2015). Scaffoldings of the affective mind. *Philosophical Psychology*, 28(8), 1157–1176. <https://doi.org/10.1080/09515089.2014.976334>.
- Condon, W. S. (1982). Cultural microrhythms. In M. Davis (Ed.), *Interaction rhythms: Periodicity in communicative behavior* (pp. 53–76). Human Sciences.
- Condon, W. S., & Sander, L. W. (1974). Neonate Movement is synchronized with adult Speech: Interactional participation and Language Acquisition. *Science*, 183(4120), 99–101. <https://doi.org/10.1126/science.183.4120.99>.
- Cordes, K., Egmore, I., Smith-Nielsen, J., Köppe, S., & Væver, M. S. (2017). Maternal touch in caregiving behavior of mothers with and without postpartum depression. *Infant Behavior and Development*, 49(May), 182–191. <https://doi.org/10.1016/j.infbeh.2017.09.006>.
- Davis, M. (Ed.). (1982). *Interaction rhythms: Periodicity in communicative behavior*. Human Sciences.
- De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences*, 6(4), 485–507. <https://doi.org/10.1007/s11097-007-9076-9>.
- De Jaegher, H., Pieper, B., Clénin, D., & Fuchs, T. (2017). Grasping intersubjectivity: An invitation to embody social interaction research. *Phenomenology and the Cognitive Sciences*, 16(3), 491–523. <https://doi.org/10.1007/s11097-016-9469-8>.
- Delafield-Butt, J. T., & Trevarthen, C. (2015). The ontogenesis of narrative: From moving to meaning. *Frontiers in Psychology*, 6 (September), 1157. <https://doi.org/10.3389/fpsyg.2015.01157>.
- Delavenne, A., Gratier, M., Devouche, E., & Apter, G. (2008). Phrasing and fragmented time in pathological mother-infant vocal interaction. *Musicae Scientiae*, 12(2008), 47–70. <https://doi.org/10.1177/1029864908012001031>.
- Devouche, E., & Gratier, M. (2001). Microanalyse du rythme. *DEVENIR*, 13(2), 55–82.
- Egmore, I., Cordes, K., Smith-Nielsen, J., Væver, M. S., & Köppe, S. (2018). Mutual regulation between infant facial affect and maternal touch in depressed and nondepressed dyads. *Infant Behavior and Development*, 50(May 2017), 274–283. <https://doi.org/10.1016/j.infbeh.2017.05.007>.
- Elmlinger, S. L., Schwade, J. A., Vollmer, L., & Goldstein, M. H. (2023). Learning how to learn from social feedback: The origins of early vocal development. *Developmental Science*, 26(2). <https://doi.org/10.1111/desc.13296>.
- Español, S., Martínez, M., & Rodríguez, F. G. (2022). An Invitation to an Embodied, Multimodal, and Intersubjective Approach to Development. In *Moving and Interacting in Infancy and Early Childhood* (pp. 1–20). Springer International Publishing. [https://doi.org/10.1007/978-3-031-08923-7\\_1](https://doi.org/10.1007/978-3-031-08923-7_1).
- Gallagher, S. (2005). *How the body shapes the mind*. Clarendon.
- Gibson, J. (1979). *The ecological approach to visual perception*. Houghton Mifflin.
- Gratier, M. (2009). Du Rythme Expressif à La narrativité dans l'échange vocal mère-bébé. *Champ Psychosomatique*, 54(2), 35. <https://doi.org/10.3917/cpsy.054.0035>.
- Gratier, M., & Apter-Danon, G. (2009). The improvised musicality of belonging: Repetition and variation in mother-infant vocal interaction. In *Communicative musicality: exploring the basis of human companionship* (pp. 301–327).
- Gratier, M., & Magnier, J. (2012). Sense and synchrony: Infant Communication and Musical Improvisation. *Intermédialités: Histoire Et Théorie Des arts. Des Lettres Et Des Techniques*, 19, 45. <https://doi.org/10.7202/1012655ar>.
- Gratier, M., & Trevarthen, C. (2007). Voice, vitality and meaning: On the shaping of the infant's utterances in willing Engagement with Culture. Comment on Bertau's on the notion of Voice. *International Journal for Dialogical Science*, 2(1), 169–181.
- Gratier, M., Devouche, E., Guellai, B., Infanti, R., Yilmaz, E., & Parlato-Oliveira, E. (2015). Early development of turn-taking in vocal interaction between mothers and infants. *Frontiers in Psychology*, 6(September), 1–10. <https://doi.org/10.3389/fpsyg.2015.01167>.
- Gratier, M., Evans, R., & Stevanovic, K. (2017). Negotiations: Sound and speech in the making of a studio recording. In E. F. Clarke, & M. Doffman (Eds.), *Distributed creativity: Collaboration and improvisation in contemporary music*. Oxford University Press.
- Hoffding, S., & Schiavio, A. (2021). Exploratory expertise and the dual intentionality of music-making. *Phenomenology and the Cognitive Sciences*, 20(5), 811–829. <https://doi.org/10.1007/s11097-019-09626-5>.

- Iyer, V. (2004). Improvisation, temporality and embodied experience. *Journal of Consciousness Studies*, 11(3–4), 159–173. <http://www.ingentaconnect.com/content/imp/jcs/2004/0000011/F0020003/art00009>.
- Iyer, V. (2016). Improvisation, action understanding, and music cognition with and without bodies. In G. Lewis, & B. Piekut (Eds.), *The Oxford handbook of critical improvisation studies* (Vol. 1, pp. 74–90). Oxford University Press.
- Jaffe, J., & Feldstein, S. (1970). *Rhythms of dialogue*. Academic.
- Jaffe, J., Stern, D. N., & Peery, J. C. (1973). Conversational coupling of gaze behavior in prelinguistic human development. *Journal of Psycholinguistic Research*, 2(4), 321–329. <https://doi.org/10.1007/BF01067054>.
- Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., Jasnow, M. D., Rochat, P., & Stern, D. N. (2001). *Rhythms of dialogue in infancy: Coordinated timing in development*. Blackwell.
- Jover, M., & Gratier, M. (2023). Toward a multimodal and continuous approach of infant-adult interactions. *Interaction Studies Social Behaviour and Communication in Biological and Artificial Systems*, 24(1), 5–47. <https://doi.org/10.1075/is.22041.jov>.
- Krueger, J. W. (2010). Doing things with music. *Phenomenology and the Cognitive Sciences*, 10(1), 1–22. <https://doi.org/10.1007/s11097-010-9152-4>.
- Krueger, J. W. (2013). Merleau-Ponty on shared emotions and the joint ownership thesis. *Continental Philosophy Review*, 46(4), 509–531. <https://doi.org/10.1007/s11007-013-9278-5>.
- Krueger, J. W. (2014a). Affordances and the musically extended mind. *Frontiers in Psychology*, 4(January), 1–13. <https://doi.org/10.3389/fpsyg.2013.01003>.
- Krueger, J. W. (2014b). Affordances and the musically extended mind. *Frontiers in Psychology*, 4(January), 1003. <https://doi.org/10.3389/fpsyg.2013.01003>.
- Krueger, J. W. (2014c). Musical manipulations and the emotionally extended mind. *Empirical Musicology Review*, 9(3), 1–5.
- Krueger, J. W. (2014d). Varieties of extended emotions. *Phenomenology and the Cognitive Sciences*, 13(4), 533–555. <https://doi.org/10.1007/s11097-014-9363-1>.
- Langer, S. K. (1953). *Feeling and form: A theory of art developed from Philosophy in a new key*. Charles Scribner's Sons.
- Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffres (Ed.), *Cerebral mechanisms in behavior: the Hixon Symposium* (pp. 112–135). Wiley. <https://doi.org/10.1016/j.humov.2007.04.001>.
- Levin, K. (2016). Aesthetic movements of embodied minds: Between Merleau-Ponty and Deleuze. *Continental Philosophy Review*, 49(2), 181–202. <https://doi.org/10.1007/s11007-016-9376-2>.
- Levin, K., Roald, T., & Funch, B. S. (2019). Visual art and the rhythm of experience. *The Journal of Aesthetics and Art Criticism*, 77(3), 281–293. <https://doi.org/10.1111/jaac.12647>.
- Maldiney, H. (2012a). Notes sur le rythme. In J.-P. Charcosset (Ed.), *Henri Maldiney: penser plus avant ... actes du colloque de Lyon (13 et 14 novembre 2010)* (pp. 17–22). Éd. de la Transparence.
- Maldiney, H. (2012b). *Regard, parole, espace* (C. Chaput, P. Grosos, & Maria. Villela-Petit, Eds.). Les Éd. du Cerf.
- Malloch, S. N. (2000). Mothers and infants and communicative musicality. *Musicae Scientiae*, 3(1999), 29–57. <https://doi.org/10.1177/10298649000030S104>.
- Malloch, S., & Trevarthen, C. (2009). Musicality: Communicating the vitality and interests of life. In *Communicative musicality: Exploring the basis of human companionship* (Issue January 2009, pp. 1–11).
- Merleau-Ponty, M. (1960). Le langage indirect et les voix du silence. *Signes* (pp. 49–104). Gallimard.
- Merleau-Ponty, M. (2007). Indirect Language and the voices of silence. In T. Toadvine, L. Lawlor, & R. McCleary (Trans.) (Eds.), *The Merleau-Ponty reader* (pp. 241–282). Northwestern University.
- Merleau-Ponty, M. (2012). *Phenomenology of perception* (D. A. Landes, Trans.). Routledge.
- Noë, A. (2012). *Varieties of Presence*. Harvard University Press.
- Pérez, S. C., & Español, S. A. (2016). Multimodal Study of Adult-Infant Interaction: A review of its origins and its current status. *Paidéia (Ribeirão Preto)*, 26(65), 377–385. <https://doi.org/10.1590/1982-43272665201613>.
- Plato. (1980). *The laws of Plato (T. L. Pangle, Trans.)*. The University of Chicago Press
- Provasi, J., Anderson, D. I., & Barbu-Roth, M. (2014). Rhythm perception, production, and synchronization during the perinatal period. *Frontiers in Psychology*, 5(SEP), 1–16. <https://doi.org/10.3389/fpsyg.2014.01048>.
- Reddy, V. (2008). *How infants know minds*. Harvard University Press.

- Rousseau, J. J. (1997). Essay on the Origin of languages [1781]. In R. Geuss, Q. Skinner, Victor, & Gourevitch (Trans.) (Eds.), *The discourses and other political writings* (pp. 247–300). Cambridge University Press.
- Schiavio, A., & Hoffding, S. (2015). Playing together without communicating? A pre-reflective and enactive account of joint musical performance. *Musicae Scientiae*, 19(4), 366–388. <https://doi.org/10.1177/1029864915593333>.
- Schiavio, A., van der Schyff, D., Cespedes-Guevara, J., & Reybrouck, M. (2016). Enacting musical emotions. sense-making, dynamic systems, and the embodied mind. *Phenomenology and the Cognitive Sciences*. <https://doi.org/10.1007/s11097-016-9477-8>.
- Sheets-Johnstone, M. (1966). *The phenomenology of dance*. The University of Wisconsin.
- Sheets-Johnstone, M. (2000). Kinetic tactile-kinesthetic bodies: Ontogenetical foundations of apprenticeship learning. *Human Studies*, 23(4), 343–370. <https://doi.org/10.1023/a:1005618313194>.
- Sheets-Johnstone, M. (2008). *The roots of morality*. Pennsylvania State University Press.
- Sheets-Johnstone, M. (2011). *The primacy of movement (expanded)*. John Benjamins Publishing Company.
- Stern, D. N. (1977). *The first relationship: Infant and mother*. Harvard University Press.
- Stern, D. N. (1985). *The Interpersonal World of the infant: A View from Psychoanalysis and Developmental psychology*. Basic Books.
- Stern, D. N. (2000). Putting time back into our considerations of infant experience: A microdiachronic view. *Infant Mental Health Journal*, 21(1–2), 21–28. [https://doi.org/10.1002/\(SICI\)1097-0355\(200001/04\)21:1/2%3C21::AID-IMHJ3%3E3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1097-0355(200001/04)21:1/2%3C21::AID-IMHJ3%3E3.0.CO;2-Z).
- Stern, D. N. (2004). *The present moment in psychotherapy and everyday life*. W.W. Norton.
- Stern, D. N. (2010). *Forms of vitality: Exploring dynamic experience in psychology, the arts, psychotherapy, and development*. Oxford University Press.
- Straus, E. (1963). In J. Needleman (Ed.), *The primary world of senses a vindication of sensory experience*. Trans.). Free Press of Glencoe.
- Straus, E. (1966). The Forms of Spatiality. In E. Eng (Trans.), *Phenomenological psychology: the selected papers of Erwin W. Straus* (pp. 3–38). Basic Books.
- Trevarthen, C. (1977). Descriptive analyses of infant communicative behaviour. In H. Rudolph, & Schaffer (Eds.), *Studies in Mother-Infant Interaction* (pp. 227–270). Academic.
- Trevarthen, C. (2000). Musicality and the intrinsic motive pulse: Evidence from human psychobiology and infant communication. *Musicae Scientiae*, 1999, 155–215. <https://doi.org/10.1177/1029864900030S109>.
- Trevarthen, C. (2004). How Infants Learn How to Mean. In Mario. Toroko & Luc. Steels (Eds.), *A learning zone of one's own: sharing representations and flow in collaborative learning environments* (pp. 37–69). IOS Press.
- Trevarthen, C. (2009). Human biochronology: On the source and functions of musicality. *Music that Works: Contributions of Biology Neurophysiology Psychology Sociology Medicine and Musicology*, 221–265. [https://doi.org/10.1007/978-3-211-75121-3\\_16](https://doi.org/10.1007/978-3-211-75121-3_16).
- Trevarthen, C. (2016). From the intrinsic motive pulse of infant actions to the Life Time of Cultural meanings. In B. Mölder, V. Arstila, & P. Øhrstrøm (Eds.), *Philosophy and psychology of Time*. Springer. <https://doi.org/10.1007/978-3-319-22195-3>.
- Tronick, E. Z., Brazelton, T. B., & Als, H. (1978). The structure of face-to-face interaction and its developmental functions. *Sign Language Studies*, 18(18), 1–16.
- Varela, F. J., Thompson, E. T., & Rosch, E. (1991). *The embodied mind: Cognitive Science and Human Experience*. The MIT Press
- Wiskus, J. (2015). *The rhythm of thought: Art, literature, and music after Merleau-Ponty*. The University of Chicago.