Fostering Social Interaction Through Sound Feedback: Sentire

MARTA RIZZONELLI, JIN HYUN KIM, PASCAL STAUDT and MARCELLO LUSSANA

Department of Musicology and Media Studies, Humboldt University of Berlin, Germany
Emails: rizzonma@hu-berlin.de; jin.hyun.kim@hu-berlin.de; pascal.staudt@hu-berlin.de; marcello.lussana@hu-berlin.de

Sentire is a body–machine interface that sonifies motor behaviour in real time and a participatory, interactive performance in which two people use their physical movements to collaboratively create sound while constantly being influenced by the results. Based on our informal observation that basal social behaviours emerge during Sentire performances, the present article investigates our principal hypothesis that Sentire can foster basic mechanisms underlying non-verbal social interaction. We illustrate how coordination serves as a crucial basic mechanism for social interaction, and consider how it is addressed by various therapeutic approaches, including therapeutic use of real-time auditory feedback. Then we argue that the implementation of Sentire may be fruitful in healthcare contexts and in promoting general well-being. We describe how the Sentire system has been developed further within the scope of the research project 'Social interaction through sound feedback–Sentire' that combines human–computer interaction, sound design and real-world research, against the background of the relationship between sound, sociality and therapy. The question concerning how interaction is facilitated through Sentire is addressed through the first results of behavioural analysis using structured observation, which allows for a quasi-quantitative sequential analysis of interactive behaviour.

1. INTRODUCTION

Sound, especially musical sound, has traditionally been viewed as a product of human activity with intrinsic sociocultural value. However, the relationship between the concepts of sound and sociality is a complex matter, raising questions regarding both the nature of musical sound and, perhaps more critically, the nature and definition of sociality. The point of departure for our discussion is a conceptualisation (proposed by some of the authors of the present article, and taking into account the musical context; Kim, Reifgerst and Rizzonelli 2019) that attempts to categorise social behaviour which comprises forms of basal social behaviour that, although not yet covered in social ontology, are nevertheless relevant to inter-individual behaviours, including those mediated by musical sound.

The categories of 'inter-individual behaviour leading to social interaction, without any collective or we-intentions', 'collective inter-individual behaviour involving we-intentions' and 'collective inter-individual behaviour based on shared or collective intentionality' (see section 2.1) indicate that sociality is the result of mechanisms that unfold on both basic and complex levels. That various salient features of music-making and listening present commonalities with the concept of sociality (Cross, Laurence and Rabinowitch 2012) suggests that musical sound may be a valid tool to facilitate social interaction and, more generally, pro-social behaviours; this has been confirmed by a number of empirical studies. Sociality and social interaction, in turn, are a fundamental and indispensable aspect of human health.

Based on the proposed relationship between sound, sociality and health, we developed a research project in which a body–machine interface and participatory performance is used to exploit the power of musical sound. This project is called 'Sentire', an Italian word meaning both 'to hear' and 'to feel'. Born as an artistic endeavour and further developed as a scientific project, Sentire offers the possibility of accessing sociality in an innovative and intuitive fashion. The goals of the project are facilitating social interaction through technologically mediated sound feedback and fostering general well-being and use in various therapeutic approaches.

To investigate these possibilities, we apply the empirical method of structured observation – well-established in the social sciences – to analyse human behaviour during social interaction with Sentire. In sections 2 and 3, we discuss the theoretical background of the project: the conditions of social behaviour and mechanisms of social interaction, with a particular focus on coordination (section 2) and the relationship between musical sound and sociality (section 3). Section 4 discusses Sentire as both a sound-based body–machine interface and a participatory performance. Finally, in section 5, we present the first results of our structured observations, and, building on the fundamental role of coordination as discussed in the theoretical background, illustrate how we have investigated Sentire's capacity to facilitate social interaction.

2. THEORETICAL BACKGROUND

2.1. Sociality: from basic to higher-ordered social behaviours

Human social interaction can occur in a variety of contexts and on different levels of complexity. More often than not, social interaction is constrained by implicit rules determined by specific social contexts, norms, or hierarchies. Social ontology – an attempt to 'explain the fundamental nature and mode of existence . . . of human social institutional reality' (Searle 2010: ix), which therefore addresses these constraints – has, however, disregarded those (largely non-verbal) forms of social interaction that do not necessarily require social cognition and intention. Kim et al. (2019: section 3) develop a more comprehensive model of social ontology, proposing distinct categories of social interaction: basic and higher-order social behaviours. Since basic social behaviours that do not involve shared or collective intentionality also play an important role in musical interaction and other forms of non-verbal interaction that do not aim to communicate representational meanings, Kim and colleagues' categories of social behaviours serve as a point of departure for our discussion of sociality and (musical) sound.

For simplicity, in outlining those categories we consider only dyadic interaction, although the conditions given are also valid for groups of more than two individuals.

First, social interaction requires two individuals that are capable of perceiving both their own and another's behaviour, and the difference between the self and the other. Second, these individuals must be situated in a shared social context that affords or constrains both the selection of an individual's behaviour and their perception of the other's behaviour. Third, both individuals have to act in a causal relation to the social context constraining them, thus creating an action–perception cycle which may or may not involve social cognition; in addition, both individuals have to act in the course of reciprocation to their counterpart.

A category of behaviours that meets all the preceding conditions is 'inter-individual behaviour leading to social interaction, without any collective or we-intentions'; this category does not require social cognition (Kim et al. 2019: section 3.2). A classic example is infant–caregiver non-verbal vocalisation. Such an interaction is neither planned nor based on intention, but is the result of both caregiver and infant fulfilling all aforementioned conditions.

The interaction may be more complex and require social cognition when individuals share what Searle terms 'we-intentions' (Searle 1990). An example of 'collective inter-individual behaviour involving we-intentions' (Kim et al. 2019: section 3.3) is continuous applause by an audience following a performance with the intent of prompting an encore. Here, the communal intent, performed individually, is oriented towards a common goal.

A higher-ordered form of social interaction is categorised as 'collective inter-individual behaviour based on shared or collective intentionality' (Kim et al. 2019: section 3.4). Language-based forms of social interaction aiming to communicate representational meanings belong to this category. Some forms of non-verbal interaction, such as orchestral performance, fall into this category as well; in such cases, each individual's act is an integral part of an overall act, and each individual adjusts or complements his own behaviour to that of the others to pursue the overall goal (e.g., playing a symphony).

Tying in with Kim and colleagues' claim that social interaction is based on the perception and selection of one's own and the other's behaviour, distinguishing the self from the other, we propose to ground social interaction on kinaesthetic perception and coordination. In particular, we want to set our primary focus on how coordination builds the foundations of basal non-verbal social behaviours. Therefore, in the next section, after briefly introducing kinaesthetic perception, we will discuss the concept of coordination in detail. Even though we recognise the importance – both for our project and for social interaction in general – of higher-ordered complex behaviours including shared and collective intentionality and possibly involving verbalisations, social norms and power dynamics, we do not address these topics in the present article.

2.2. Coordination as a basic mechanism for social interaction

One of the necessary conditions for coordination is kinaesthetic perception, that is, the perception of one's own movements (Rosenbaum 2010: 51). According to some conceptualisations, kinaesthetic perception can also be oriented towards a partner's movements (see 'kinaesthetic intersubjectivity', Samaritter and Payne 2013 and 'enkinaesthesia', Stuart 2012) and thus build a bridge towards social interaction.

Coordination is a broad concept used to encompass a variety of behaviours. We propose to distinguish two main acceptations: that referring to the temporal axis, where 'coordination' can be replaced by the more specific terms of 'synchronisation' and 'entrainment' (generally used interchangeably); and that extending beyond the temporal axis and referring to the matching of specific behavioural aspects, which we will call 'attunement'.

Coordination as synchronisation/entrainment is defined as 'an adjustment of rhythms of oscillating objects due to their weak interaction' (Pikovsky, Rosenblum and Kurths 2001: 8).

The concept of synchronisation/entrainment is established as a physical phenomenon but it has been taken up and further developed in the social sciences (McGrath and Kelly 1986) and other disciplines, including musicology (Clayton, Sager and Will 2005; Clayton, Jakubowski and Eerola 2019; Kim et al. 2019).

Many empirical studies show that synchronisation/entrainment facilitates or strengthens pro-social behaviour in the form of empathy and affiliation. Coordination within long-term musical group interaction was shown to increase the index of empathy and the ability to recognise emotional facial expressions (Rabinowitch, Cross and Burnard 2012). Stupacher (Stupacher, Wood and Witte 2017) found that synchronisation/entrainment as performed in a tapping task had a positive effect on an implicit measure of affiliation and helpfulness.

A theoretical model of the mutual connection between coordination and social interaction has been provided by De Jaegher and Di Paolo (2007). Based on the premises that 1) the ability to create and appreciate meaning, which is called sense-making, is the definitional[1] property of cognisers (ibid.: 488), and 2) that movements (including utterances) are the tools of our cognition, the authors propose participatory sense-making as the core of social cognition. This conceptualisation emphasises the temporal aspects of interaction, particularly coordination, as borrowed from dynamical systems theory. In line with our views, De Jaegher and Di Paolo understand coordination as synchronisation/entrainment and define social interaction as the coupling, regulated by coordination, of at least two autonomous agents. Coordination allows the agents to maintain their coupling while remaining autonomous and enables (temporal) organisation to emerge from the social relation of the agents. Depending on the agents' participation in the creation of meaning, participatory sense-making may be weaker (interaction is coordinated but coordination largely remains an individual activity) or stronger (interaction is the result of a fully joint and shared effort).

The model can be summarised as follows: movements (and utterances) are the tools of sense-making, while regulation of social coupling takes place through coordination of movements (and utterances); therefore, social agents can coordinate their sense-making in social interaction. Coordination influences (sustains, modifies) interaction and interaction in turn promotes certain patterns of coordination. If coordination is the crucial feature of participatory sense-making and participatory sense-making defines social cognition, then coordination represents the basis on which other, more complex forms of social interaction and cognition are built.

When participatory sense-making is combined with mutual incorporation – or, as we suggest, mutual embodiment – that is, the process by which two people extend the perception of their bodies and form a common intercorporality, enactive intersubjectivity emerges: the ability 'to grasp (to a certain extent) the experience of a non-verbal partner by interacting with him' (Fuchs and De Jaegher 2009: 482).

We now turn to the conceptualisation of coordination as attunement. Coordination can be understood in a much broader sense than simply the adjustment of two behaviours to each other on the temporal axis. In fact, rather than the matching of another person's behaviour per se, it can indicate the matching of a specific *aspect* of that behaviour. Coordination therefore need not be an exact reproduction of the original, but can be a response that resembles the original only in certain respects. This form of coordination has been observed in infant–caregiver interaction from approximately nine months old: Stern (1985) reports the case of a boy sitting in front of his mother and shaking a rattle up and down with a display of interest; the mother's reaction is that of nodding her head up and down to the child's beat. Mother and son are not moving the same body part, yet the scene gives the impression that a form of matching has occurred. Stern defines this form of matching 'affect attunement', in which one party matches a behavioural aspect that reflects the feeling state of their partner. Indeed, in mother–infant dyads, affect attunement is a far more common maternal response than pure imitation (48 per cent versus 19 per cent, respectively).

Furthermore, in 87 per cent of cases affect attunement occurs cross-modally, meaning that the two interactors are using different modalities,[2] and transferring so-called amodal properties (intensity, time, shape) from one modality to another (Stern, Hofer, Haft, and Dore 1985). The ability to identify cross-modal equivalences arises in the first months of life and grows as a distinctive property of human interaction. Interaction between adults is also rich in cross-modal attunement experiences such as smiling at a known person that approaches us; for example, in a work context.

Although the concept of affect attunement seems very promising for the analysis of interaction, it has not been consistently integrated into sociological

---

[1] 'The distinction between a strictly physical encounter and a cognitive one is to be found in the dimension of significance for the cogniser itself that is characteristic only of the latter class, even though cognitive interactions are themselves also physical processes' (De Jaegher and Di Paolo 2007: 488).

[2] The most common matching occurs between the modalities of body movements and vocality. This type of cross-modal interaction cannot be understood as verbal in the strict sense, as vocality is involved, but verbalisation or semantic meaning is not.

research since its introduction in the mid-1980s, nor is there another established term to describe this specific form of coordination. As such, we have chosen to use 'attunement', without any further specifications, to indicate coordination in social contexts that refers to the matching of a specific behavioural aspect and differs from synchronisation/entrainment.

3. MUSICAL SOUND AND SOCIALITY

3.1. Sociality of musical behaviours

The role of coordination in social interaction has come to the fore in recent research on proto-musical and proto-linguistic behaviours. In developmental psychology, the coordination that emerges during infant–caregiver interaction is considered to be based on 'communicative musicality', an ability linked to 'our innate skill for moving, remembering and planning in sympathy with others that makes our appreciation and production of an endless variety of dramatic temporal narratives possible' (Malloch and Trevarthen 2009: 4). The concept of communicative musicality allows verbal or musical interaction to be founded on pre-linguistic interactions using sound and gesture. Hence, to discuss the social dimensions of sound, we are interested in sonorous phenomena and behaviours that become coordinated during interaction.

Sonorous phenomena and behaviours that become coordinated with each other, or with gestural phenomena and behaviours, may be characterised as forms of movement that have duration and contours of intensity, pitch and so on. Coordination may be related to the structure of duration (i.e., rhythmic structure) as synchronisation/entrainment; and/or to the structures of contours as attunement. Even sonorous phenomena and behaviours used in verbal communication often become coordinated with each other when interacting speakers make sense of and empathise with each other through tacit engagement (Gill 2015).

Musical sounds relate to one another to build dynamic forms of movement, which can be referred to as 'sonorous forms of vitality'. Stern coined the term 'forms of vitality' to describe physical actions and mental processes that are shaped in relation to others and allow for the experience of vitality (Stern 2010). Forms of vitality in music manifest themselves in structural features (e.g., contours of pitch) and codified forms (e.g., sonata form; Stern 2010, Kim 2013). This argument is in line with several music-theoretical approaches characterised as 'energetic' (Rothfarb 2002), in which music's dynamic qualities are identified and described in reference to contours of force (Mersmann 1925; Kurth 1931; Zuckerkandl 1956; Halm 1978). Since forms of vitality emerge relationally, that is, through interaction with others, forms of vitality in music that resemble human forms of

vitality could act as a basis for investigating the social dimensions of musical sounds.

Our thesis holds that music's structural features and codified musical forms constitute sonorous forms of vitality that are related both to musical elements and to the world and others. We thereby refute the view of musical formalism in aesthetics, which sees 'tonally moving forms' as constituted merely through intra-musical relationships and not through relationships with the world and others (Hanslick 1854).

Furthermore, music-making often comprises acts of collaboration, which involve either we-intentions or collective intentionality (see section 2.1). At the same time, listening to music involves mental processes of co-shaping forms of vitality (Kim 2013), which could correspond to those forms of vitality that manifest themselves in musicians' mental processes, as well as music's structural features and codified musical forms. Hence, music affords a sense of mutual affiliation involving coordination processes among musicians, or between musicians and listeners (Cross 2014). These processes include synchronisation/entrainment related to movements, tempo and metrical structure (Clayton et al. 2019), as well as processes relevant to joint action (Keller, Novembre and Hove 2014) that can be subsumed under the concept of attunement.

Cross argues that music plays a significant social role at times of potential social stress (Cross 2009; Cross and Woodruff 2009). He finds that music acts as a medium for managing events that can be categorised as situations of social uncertainty (Cross 2014) by facilitating reciprocity, social bonding and cohesion (Cross 2007; Woolhouse, Tidhar and Cross 2016). Cross proposes the term 'empathic creativity' to indicate the mutual affective alignment that arises during the creative process of making music together (Cross et al. 2012: 6). Empathic creativity indicates both the innate mechanisms that occur on an automatic level below awareness and the acquired processes that occur on a volitional, aware level. Musical interaction exhibits various salient features that may give rise to empathic creativity in cooperative musical contexts: imitation, synchronisation/entrainment, disinterested pleasure, flexibility, ambiguity and shared intentionality (for a detailed description, see Cross et al. 2012).

Empirical studies (Cross et al. 2012) have demonstrated that the aforementioned natural features of musical interaction can promote and train empathic creativity. Digital musical instruments can be specifically designed for such purposes, allowing for a collective experience and guaranteeing intuitiveness of the system and easy access for all users (Blaine and Fels [2003] 2017). A central feature of these instruments is limitation of the musical parameters that can be controlled (Cook [2001] 2017; Tanaka and Knapp

2002), so that interaction is prioritised (Robson 2002) and made possible regardless of users' musical skills.

3.2. Musical sound, therapy and well-being

Sound, particularly in the form of concurrent (or real-time) auditory feedback, can be used to support kinaesthetic perception (briefly introduced in section 2.2). The potential of concurrent sensory feedback for enhancing motor learning and performance has been extensively studied (see Sigrist, Rauter, Riener and Wolf 2013, and Effenberg, Fehse, Schmitz, Krueger and Mechling 2016, respectively). For instance, the relationship between kinaesthetic perception, sound and intersubjectivity has been explored in dance improvisation. Here, the individual's action–perception loop is combined with the loop between the self and the other in a form of non-representational, embodied and experiential interaction that Samaritter defines as 'kinaesthetic intersubjectivity' (Samaritter and Payne 2013). The therapeutic approach of dance movement psychotherapy is based on this principle and makes use of shared movement as a specific intervention.

In the therapeutic context, concurrent auditory feedback has been applied (e.g., for patients with Parkinson's disease) to improve specific gait parameters such as stride time (Hove, Suzuki, Uchitomi, Orimo and Miyake 2012) and stride length (Rizzonelli, Kim, Gladow and Mainka 2017). In both study designs, interactive feedback is compared with non-interactive auditory stimulation and no stimulation. The results of both studies show that the condition under research corresponds to significant improvement in the training's goal parameter as compared with the two control conditions. The assumptions that underlie these therapeutic protocols are: 1) that musical feedback can close the disrupted action–perception loop that characterises the impaired proprioception typical of Parkinson's disease; and 2) that mutual entrainment between motor execution and auditory perception is the mechanism that enables the closure of the loop.

It may be argued that the preceding studies show the effectiveness of real-time auditory feedback on motor learning or task performance, but do not support free improvisational movement as a form of non-verbal social interaction. However, we claim that, because concurrent auditory feedback has been proven effective in motor tasks that rely on the basic mechanism of synchronisation/entrainment, it is likely to be effective even in the context of social interaction, where both basic and more complex cognitive processes are at play. Moreover, in line with several studies suggesting that, with respect to motor performance, music is a more effective stimulus than isolated rhythmic beats (Thaut, McIntosh, Rice, Miller, Rathbun and Brault 1996; Thaut, Rathbun and Miller 1997; Dyer, Stapleton and Rodger 2017; Rose, Delevoye-Turrell, Ott, Annett and Lovatt 2019), we propose the use of *musically* informed real-time feedback.

Recent musicological research has considered which musical parameters influence motor production (e.g., Buhmann, Desmet, Moens, Van Dyck and Leman 2016) and encouraged further research on 'the precise aspects of music, besides musical tempo, that might influence the spatialisation of body movements' (Styns, Van Noorden, Moelants and Leman 2007). This leads to a reflection on how sound can be shaped involving musicality (i.e., with regard not only to timing but also to pitch, dynamics and so on) to facilitate therapy and, more generally, foster well-being. That many studies applying sound with therapeutic purposes do not describe the acoustic stimulation in detail reveals that the choice of the stimulation and its intrinsic qualities are often considered of secondary importance with respect to the goal of the studies. By contrast, we believe that the musicality available in sound should be central for the development of any sound-based intervention – be it therapeutic or aimed at increasing general well-being.

Sound has been described with great accuracy in a project aimed at promoting reflective engagement with the act of walking through acoustic feedback. Feltham (Feltham, Loke, van den Hoven, Hannam and Bongers 2014) developed an interactive surface called 'Slow Floor' that generates sound according to the pressure exerted by the person walking on it. Pressure changes are mapped to sound qualities such as pitch and volume that seek to resemble the qualities of the movements, providing an intuitive experience. Each sound environment used in the study includes both pitched and unpitched acoustic material. Feltham's project combines the facilitation and kinaesthetic perception with sound design and human–computer interaction. First empirical findings showed that users experienced a strong sense of creative agency over the sound, as they were stimulated to create new, original foot movements while generating and responding to the sound at the same time.

In a similar project by Françoise and colleagues (Françoise, Candau, Alaoui and Schiphorst 2017), the interactive system called 'still, moving' was developed to enhance kinaesthetic perception[3] through real-time sonification of a person's micro-movements. The system consists of two bracelets worn on the user's legs, which track movement information and muscle tension through electromyographic sensors, with the

---

[3]The authors refer to 'kinaesthetic awareness' but define it as 'the perception of our position and movement in space'. Therefore, for the sake of consistency with our previous argumentation, we use the term 'kinaesthetic perception'.

system's sensitivity increasing as the extent of movement decreases. The only sound parameter mapped to movement force is loudness, but the sound corpus is built in such a way that loudness variations result in consistent and rich timbric variations. The authors chose environmental sounds because of their evocative power, which resonates with people's auditory experiences in the real world and thus facilitates an embodied experience.

A final remark needs to be made addressing personal musical taste. Although we acknowledge that preferences and cultural background influence a person's reaction to musical stimuli – and are accordingly planning to consider personal musical taste in further studies – given the complexity and variety of the topic, our first musical feedback design, which we describe in this paper, focuses on functional, rather than subjectively perceived, acoustic qualities.

4. SENTIRE

The reflections of Feltham et al. (2014) and Françoise et al. (2017) are particularly relevant to our project. Sentire was initiated with the claim that an artistically sensitive and responsive use of sound in the form of auditory feedback can facilitate therapy and promote well-being by fostering the experience of the self as dependent on others. Building on knowledge gained from years of performance practice and on the theoretical background discussed in section 2, our project presents a developing scientific approach to adapt an artistic endeavour for healthcare purposes.

4.1. Sentire as a body–machine interface

A body–machine interface (BMI) is a technological interface capable of extending or replacing human capabilities. Utilising a BMI, the user may gain complete or shared control over the machine through signals derived from their body (Casadio, Ranganathan and Mussa-Ivaldi 2012). Sentire is a sound-based BMI that sonifies motor behaviour in real time by detecting proximity and touch between two users. While established systems capable of distance detection typically use infrared or ultrasonic sensors, Sentire uses the body itself as an electronic component of the sensor system. This overcomes the disadvantages of measuring at discrete points and/or requiring a clear field of vision.

Sentire enables whole-body proximity detection independent of sensor positioning and orientation of the body in space. This is accomplished with our custom software and capacitive sensing system. The technical implementation of proximity detection takes advantage of the human body's electric conductivity. Through a cable connected to a conducting bracelet, an electrical signal of very low voltage is fed into a

person (for convenience called the transmitter). Attaching a second person (the receiver) to the same electrical circuit with another bracelet and cable causes capacitive coupling between the two bodies. Through a low noise signal amplifier, the amplitude of the transmitted signal is measured at the receiver. Because the signal strength is dependent on the distance between the receiver and the transmitter, this 'proximity effect' can be utilised to obtain a control signal, which changes linearly with the distance of the two bodies. The signal is then mapped to selected parameters of an algorithmic sound synthesis environment (parameter mapping). Finally, the output of the sound synthesis process is made audible, enabling closed-loop auditory interaction between the two persons. Figure 1 illustrates the signal path and processing of Sentire's various components.

The use of parameter mapping is a central element in the design process of the interactive system, as it manifests the linkage between non-verbal social interaction and auditory feedback. Parameter mapping consists of the specific correspondence between control parameters (derived from performers' actions) and sound synthesis parameters (Hunt, Wanderley and Paradis 2002). As with digital musical instruments, the mapping layer is a key factor in determining constraints (Magnusson 2010), dimensionality (Gurevich and von Muehlen 2001; Zappi and McPherson 2015), and expressiveness (Arfib, Couturier and Kessous 2005). Following Rovan's classic definition of parameter mappings (Rovan, Wanderley, Dubnov and Depalle 1997), Sentire uses a 'divergent' mapping, where a one-dimensional gestural parameter (proximity) is simultaneously linked to multiple musical parameters.

Specifically, proximity is sonified in the intimate (0–0.5 m), personal (0.5–1.2 m) and social (1.2–3 m) areas, as they are defined in terms of proxemics (see Hall 1966). Depending on context or conditions, the touch feature and the corresponding intimate area can be silenced. The physical experience of proximity (and touch) within dyadic interaction is thus digitally designed as a sonic-aesthetic experience (for an overview of the concept of digital proxemics, see McArthur 2016). The design of such an experience depends strongly on multisensory integration, the automatic process through which the human nervous system integrates different sensory modalities. In comparison with unimodal events, multisensory events can be detected faster and more easily, and improve detection sensitivity and event comprehension, as more sensory information is available (Hobeika 2017). The usage of Sentire therefore combines the perceptive modalities of hearing, touch and kinaesthetic perception.

Sentire provides two sound environments (SEs; see Livingstone and Miranda 2004) called 'Sinus' and
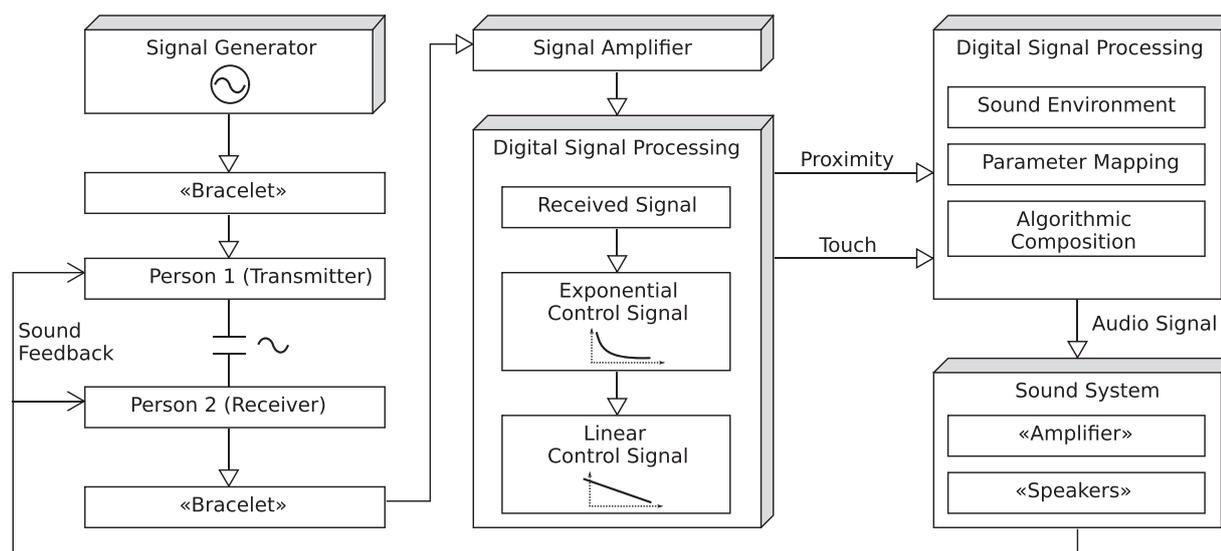
Figure 1. Signal path and processing of Sentire's various components.

'Pulse'. We understand an SE as a set of parameters that defines the tonal quality and results from predefined artistic decisions combined with a specific parameter mapping. The composition of a given SE is based on sound design (i.e., how the sounds are generated, structured, transformed and mixed) and algorithmic processes (which are optionally controlled through the parameter mapping).

The two SEs primarily differ in their tonal quality (an ambient pad sound for 'Sinus' and a percussive sound for 'Pulse') but both feature a similar mapping of the control signal to the musical parameters, which is aimed at sonically intensifying the acts of approach and touch. The proximity signal is mapped to amplitude and pitch for each sound; when the users approach each other, the sound becomes louder and higher in pitch. (In 'Pulse', the speed of the discrete pulses is also mapped to the proximity signal, that is, the closer the participants are, the faster the pulse sounds are generated.) When the users touch each other, the root frequency of the proximity sound is changed based on a given probability distribution. In addition, another synthesiser is triggered; this synthesiser has the same sound generation as that mapped to the proximity parameter, but uses an envelope with short attack and release times, which makes the touch sound percussive and thus emphasises the touch event.

4.2. Sentire as a participatory performance

Sentire is not only a body–machine interface but also a participatory performance, in which a performer invites one person at a time to interact on stage. The performer, who knows Sentire and its sound environments, applies strategies to shape the interaction.

In this sense, the performer serves as the leader of the interaction.

Through numerous performances in festivals and events worldwide,[4] we developed different strategies to intensify the interactive experience. Many of these rely on the performer's improvisational skills and are influenced by dance practices such as contact improvisation. For example, moving in a circle is a strategy to overcome the tension that may arise when expert and participant first face each other but have not yet started interaction: the expert moves along an imaginary circle on the floor and approaches the participant from the side, rather than from the front.

Through a three-year project funded by the German Federal Ministry of Education and Research (BMBF) and conducted at the Humboldt University of Berlin, we are investigating the potential of Sentire in promoting social interaction in both real-world contexts and specifically designed experiments, both for the general population and among people with specific therapeutic needs. The results of these studies will inform further technical development as well as therapeutic approaches.

5. AN EMPIRICAL APPROACH THROUGH STRUCTURED OBSERVATION

5.1. Structured observation for behavioural analysis

Observational measurements are particularly effective for the analysis of complex non-verbal behaviour, which unfolds in time and is context-sensitive, because they allow for an observation of spontaneous – rather

---

[4]For a full list, see https://sentire.me (accessed 1 December 2021).

than artificially elicited – behaviour, which is central in real-world research (Robson and McCartan 2016).

Structured observation (SO) consists of human, qualitative, non-participatory observation and codification of behaviour resulting in quantitative data such as number, frequency and duration of the codes applied. SO includes two interdependent processes: 1) segmentation of behaviour into chunks and 2) annotation, that is attribution of a code to each chunk. While segmentation of verbal communication is intuitive due to the presence of clearly defined units of speech (phonemes, words, sentences), chunking the seamless flow of non-verbal behaviour may be extremely challenging and is more prone to subjective judgement. The possibility to segment non-verbal behaviour depends entirely on the development of clear, exhaustive and mutually exclusive codes (Bakeman and Quera 2011), that is, on the elaboration of a solid coding scheme and manual. Codes can be either empirical or functional. Empirical codes are as concrete as possible and based on pure observation of the morphology (e.g., 'moves arm upwards') or the purpose (e.g., 'grasps food') of the observed action. Functional codes are more abstract and require a certain degree of inference; for example, for the observation of an infant crying, the functional code could be 'seeks attention' or 'expresses discomfort' (Bakeman and Quera 2011: 19).

To investigate original research questions, it is often necessary to develop a coding scheme *ex novo* during a complex and time-consuming initial phase of informal observation. Once scheme and manual have been refined and tested, SO in the strict sense can be applied. Cohen's kappa (Cohen 1960) is used to calculate inter-observer agreement (IOA) between two independent observers for at least 20 per cent of the data (see, e.g., Pellecchia, Beidas, Mandell, Cannuscio, Dunst and Stahmer 2020). IOA values that can be considered strong (see McHugh 2012) prove the validity of a coding scheme and the reliability of the method *in toto*.

5.2. Structured observation for the analysis of Sentire

The SO approach is ideal for the study of interaction (Fuchs and De Jaegher 2009), specifically human–computer interaction and human–robot interaction (Seifert and Kim 2008; Kim et al. 2010). Our main hypothesis, that is, that Sentire can foster non-verbal social interaction, is addressed by SO as follows: 1) what are relevant social behaviours that emerge from interaction with Sentire? 2) does an increasing trend of pro-social or interaction-relevant behaviours emerge when Sentire is used over a span of time?

To apply SO, videos are recorded by three cameras in a triangular setup (to guarantee an accurate capture of the full interaction space) and analysed with the open source software ELAN from the Max Planck Institute for Psycholinguistics. In the following, we present a detailed report of our informal observation, which has led to the development of an effective coding scheme (with strong IOA values) and to preliminary results on the interaction itself.

We started our informal observation with the annotation of all visible behaviours, then used an affinity diagram to group them into 14 categories (i.e., 14 tiers or levels of analysis in ELAN), each with a predefined set of applicable codes (controlled vocabulary). Some major problems arose during this phase: first, the definitions on which codes were based (Koch 2014, based on Kestenberg Movement Profile, see Kestenberg and Sossin 1973) were often difficult to apply in a consistent, non-interpretative way; second, some behaviours needed to be coded on multiple levels, which makes the coding process more error-prone.

We hypothesised that reducing the scheme to its minimal terms, that is, to a single category/tier of mutually exclusive and exhaustive codes, would have allowed us to overcome the aforementioned difficulties and to perform a faster and easily comparable analysis. Under the umbrella category of behaviour, we subsumed the following five codes. The codes 'simultaneous copy' and 'lagged copy' indicate coordinated behaviour understood as synchronisation/entrainment. This is based on the fact that interpersonal coordination in the form of synchronisation/entrainment, like synchronisation/entrainment in the physical domain, refers to temporal coordination both in a one-to-one ratio and in any other constant phase relationship, that is, including a time lag. The code 'compensation' accounts for behaviour that does not appear as a copy but can nonetheless be conceived as coordinated in the sense of attunement (see section 2.2) – in our case: step forward–step backwards. Finally, the codes 'different, directed behaviour' and 'different, non-directed behaviour' indicate non-coordinated or divergent behaviour (Burgoon, Dillman and Stern 1993), where head and gaze may or may not be directed towards the partner. All codes are annotated only if they last at least three seconds, which represents the minimum duration for most cognitive activities and is characterised as psychological present (see, e.g., Jaffe and Feldstein 1970; Wittmann and Pöppel 1999/2000; and Stern 2004 for a comprehensive, psychological approach).

The preceding coding scheme was systematically applied by three independent observers to a pilot study where two participants with a diagnosis of psychosomatic disorders interacted with a trained music therapist over three sessions. The analysis revealed two problems. First, it was necessary to separate segmentation from annotation so that an expert observer pre-segmented the video material before two external

Table 1. Tiers applied to each interactor individually

Whole_body	Arms	Touch	Gaze_at_other	Smile

no_steps
  step/s_forward
  step/s_backwards
  step/s_laterally_R
  step/s_laterally_L
  circle_clockwise
  circle_counter_clockwise
  turning_steps
  turning_no_steps
	rest_position
	  R_towards_partner
	  L_towards_partner
	  Both_towards_partner
	  R_not_towards
	  L_not_towards
	  Both_not_towards
		R_HAND_momentary
		L_HAND_momentary
		Both_HANDS_momentary
		R_HAND_repetitive
		L_HAND_repetitive
		Both_HANDS_repetitive
		R_HAND_sustained
		L_HAND_sustained
		Both_HANDS_sustained
		other_body_parts
			Annotated only when gaze towards the partner's face occurs.	*Annotated only when a smile occurs.*

observers could annotate. Second, despite consistent segmentation and strict formulation of the codes, IOA was insufficient. Cohen's kappa value ($0 \leq k \leq 1$; Cohen 1960) reached a mean value of 0.43, which, according to McHugh's interpretation (McHugh 2012), can only be considered weak. The reason for the ineffectiveness of the single-tier strategy probably lies in the fact that, in order for the codes to be as few and exhaustive as possible, they were not strictly empirical, but included a certain degree of interpretation.

Therefore, it was necessary to resort to the original, more time-consuming strategy with multiple tiers (effectively applied in similar studies, e.g. Evola, Skubisz and Fernandes 2015). This time, however, we maintained codes strictly empirical, did not set a minimum duration for behaviour to be coded and limited the options of the controlled vocabulary. As a result, observers were able to perform segmentation and annotation at once, that is, without pre-segmentation, and to reach a mean Cohen's kappa value of 0.85 (Cohen 1960), corresponding to strong IOA (McHugh 2012). Table 1 shows the tiers applied to each interactor individually (during the COVID-19 pandemic, we carried out studies without touch to comply with COVID-19 regulations and therefore not using the 'touch' tier).

This coding scheme is strictly and intentionally limited to empirical codes, which are not prone to subjective interpretation and can therefore be annotated consistently by independent observers. Higher-order social categories are interpreted strictly as a result of empirical observations. Single empirical codes referring to whole-body movements do not disclose information about the amount or quality of the interaction, but the combination of participant and therapist's corresponding tiers allows us to develop functional categories out of the originally annotated empirical tiers. In other words, coordination as synchronisation/entrainment results from the overlapping of the two participants' whole-body tiers. Empirical codes referring to partner-oriented arm movements,

all touch-related codes (currently not being used), and the annotation of gaze and smile, do as such provide information about directedness towards the partner and can be interpreted as attunement in Stern's sense (see section 2.2). Moreover, mutual gaze or smile result from the overlapping of the two participants' tiers 'gaze' and 'smile', respectively.

Figure 2 shows the percentage of interaction-relevant behaviours performed by the two participants during the whole pilot study. Figure 3 shows the increasing percentage of mutual gaze and smile (between each participant and the therapist) over three sessions.

The preliminary IOA results of the pilot study show the reliability of the coding scheme, which can now be applied to further studies, checking for IOA only for a portion of the data (at least 20 per cent, as indicated in the previous section). As it has become clear, the development of the current coding scheme is the result of a reiterative process, where literature-based hypotheses and codes had to be applied empirically and could be accepted only if they led to sufficient IOA. As for the results of the pilot study with respect to interaction-relevant behaviours, our preliminary findings suggest that pro-social behaviours such as coordination and directedness are solidly present during the usage of Sentire. In particular, mutual gaze and smile between each participant and the music therapist have increased over the three sessions. Clearly, these results are not suitable to claim any statistical significance, but represent a valid basis for the design of further studies and for a consistent application of SO based on the last version of the multiple-tier coding scheme.

6. CONCLUSION

The ongoing research project 'Social interaction through sound feedback – Sentire' aims to further develop Sentire, an innovative sound-based body–machine interface and participatory performance in order to promote non-verbal social interaction (both artistically guided and otherwise), specifically for therapeutic purposes. It
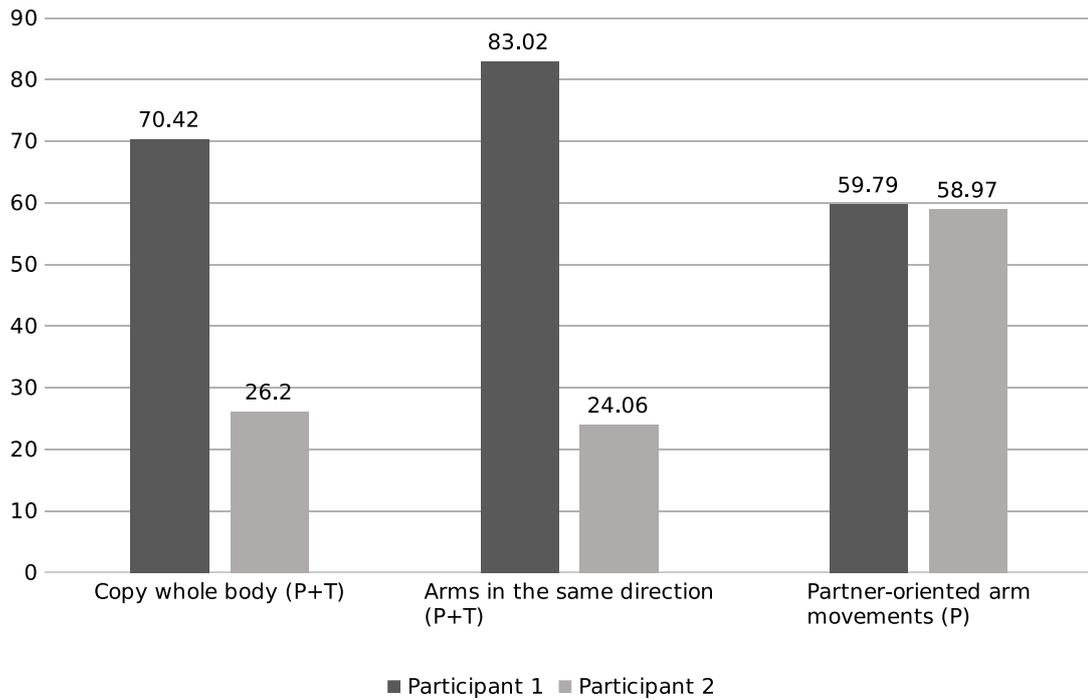
Figure 2. Percentage of interaction-relevant behaviours performed by the two participants during the whole pilot study (P = participant; T = therapist).
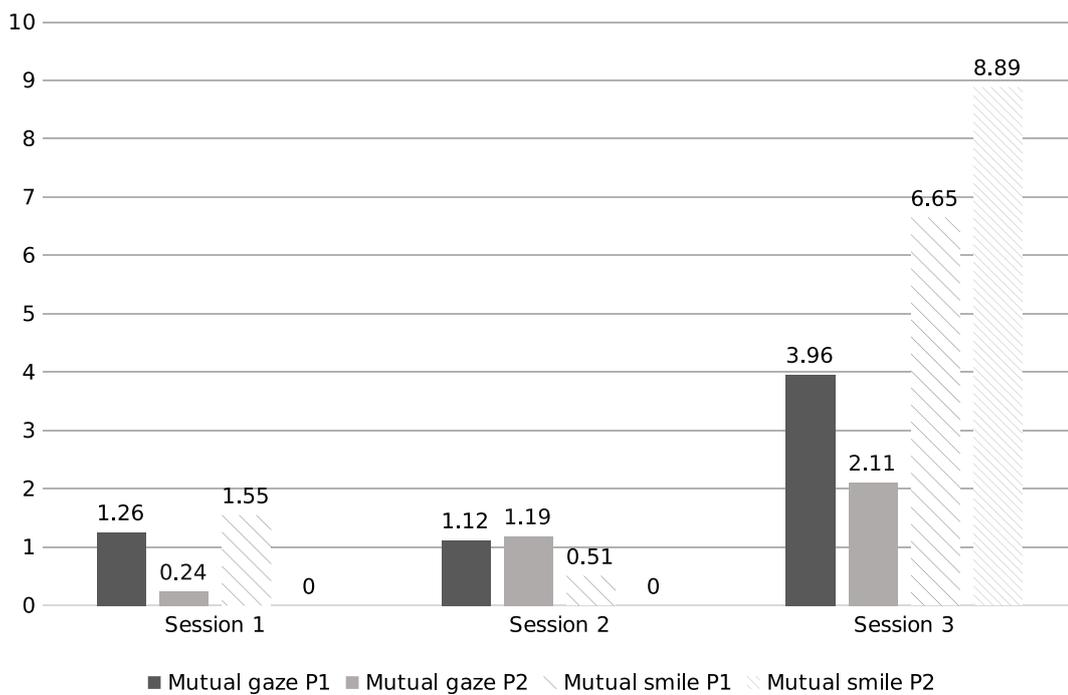
Figure 3. Increasing percentage of mutual gaze and smile (between each participant and the therapist) over three sessions (P = participant).

seeks to design the experience of social relations between interactors and the emergence of the self as dependent on others using sound feedback to interactive behaviours. The artistically informed approach is combined with human–computer interaction research and real-world research.

Tying in with considerations by Kim et al. (2019) concerning relevant conditions and systematic categories of social behaviour, we presented our theoretical claim that coordination represents a crucial basic mechanism of social interaction (sections 2.1 and 2.2). We argued that musical sound can promote sociality in the sense of communicative musicality, experience of vitality, and musical empathic creativity (section 3.1). On these grounds, we proposed that sociality facilitated through real-time auditory feedback can play a relevant part in therapeutic and well-being-related applications, emphasising that auditory feedback has the potential to increase embodied experiences and coordination (section 3.2).

The Sentire project, combining human–computer interaction with artistic and empirical approaches, was presented (section 4). The development of the interactive technology and choice of its sound design was largely derived from the artistic background of the project's creators. Artists are often pioneers in exploring and extending the use of emerging technologies; Sentire demonstrates how an art project can inform academic research, providing a technology that is both intuitive to use and fulfils all requirements for promoting sociality through musical interaction and mutual embodiment. Scientific research demands a solid methodology to empirically prove what has been shown informally in the artistic context; to this end, the potential of Sentire to foster non-verbal social interaction is investigated by applying structured observation, a quasi-quantitative method of behavioural analysis. Although this approach is suitable for the study of interaction, chunking the seamless flow of non-verbal behaviour is not a trivial matter and is prone to subjective judgement; the reiterative process of developing a coding scheme and the results of the current scheme, which shows strong levels of IOA and therefore reliability, were discussed (section 5).

Artistic and technological ideas for Sentire will be developed further, informed by the results of empirical studies; and vice versa, the design of empirical studies will be informed by artistic realisations and technological developments. We aim to establish this unique approach, Artistic Human–Computer Interaction Design, for use in diverse fields of artistic and scientific research and to explore its possible therapeutic applications.

Acknowledgements

REFERENCES

Arfib, D., Couturier, J. M. and Kessous, L. 2005. Expressiveness and Digital Musical Instrument Design. *Journal of New Music Research* **34**(1): 125–36.

Bakeman, R. and Quera, V. 2011. *Sequential Analysis and Observational Methods for the Behavioral Sciences*. New York: Cambridge University Press.

Blaine, T. and Fels, S. [2003] 2017. Contexts of Collaborative Musical Experiences. In A. Jensenius, M. Lyons (eds.) A NIME Reader. *Current Research in Systematic Musicology*, vol. **3**. Cham: Springer, 71–87.

Buhmann, J., Desmet, F., Moens, B., Van Dyck, E. and Leman, M. 2016. Spontaneous Velocity Effect of Musical Expression on Self-paced Walking. *PLoS ONE* **11**(5): e0154414. https://doi.org/10.1371/journal.pone.0154414.

Burgoon, J. K., Dillman, L. and Stern, L. A. 1993. Adaptation in Dyadic Interaction: Defining and Operationalizing Patterns of Reciprocity and Compensation. *Communication Theory* **3**(4): 295–316.

Casadio, M., Ranganathan, R. and Mussa-Ivaldi, F. A. 2012. The Body-machine Interface: A New Perspective on an Old Theme. *Journal of Motor Behavior* **44**(6): 419–33.

Clayton, M., Jakubowski, K. and Eerola, T. 2019. Interpersonal Entrainment in Indian Instrumental Music Performance: Synchronization and Movement Coordination Relate to Tempo, Dynamics, Metrical and Cadential structure. *Musicae Scientiae* **23**(3): 304–31.

Clayton, M., Sager, R. and Will, U. 2005. In Time with the Music: The Concept of Entrainment and Its Significance for Ethnomusicology. *European Meetings in Ethnomusicology* **11**: 3–142.

Cohen, J. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* **20**(1): 37–46.

Cook, P. [2001] 2017. Principles for Designing Computer Music Controllers. In A. Jensenius, M. Lyons (eds.) A NIME Reader. *Current Research in Systematic Musicology*, vol. **3**. Cham: Springer, 1–13.

Cross, I. 2007. Music and Cognitive Evolution. In L. Barrett and R. Dunbar (eds.) *The Oxford Handbook of Evolutionary Psychology*. Oxford: Oxford University Press, 649–67.

Cross, I. 2009. The Evolutionary Nature of Musical Meaning. *Musicae Scientiae* **13**(2): 179–200.

Cross, I. 2014. Music and Communication in Music Psychology. *Psychology of Music* **42**(6): 809–19.

Cross, I., Laurence, F. and Rabinowitch, T. 2012. Empathy and Creativity in Group Musical Practices: Towards a Concept of Empathic Creativity. In G. E. McPherson, G. F. Welch (eds.) *The Oxford Handbook of Music Educationi*, vol. **2**, 2nd edn. New York: Oxford University Press, 337–54.

Cross, I. and Woodruff, G. E. 2009. Music as a Communicative Medium. In R. Botha and C. Knight (eds.) *The Prehistory of Language*, vol. **1**. Oxford: Oxford University Press, 113–44.

De Jaegher, H. and Di Paolo, E. 2007. Participatory Sense-making: An Enactive Approach to Social Cognition. *Phenomenology and the Cognitive Sciences* **6**(4): 485–507.

Dyer, J., Stapleton, P. and Rodger, M. 2017. Advantages of Melodic over Rhythmic Movement Sonification in Bimanual Motor Skill Learning. *Experimental Brain Research* **235**: 3129–40.

Effenberg, A. O., Fehse, U., Schmitz, G., Krueger, B. and Mechling, H. 2016. Movement Sonification: Effects on Motor Learning Beyond Rhythmic Adjustments. *Frontiers in Neuroscience* **10**(219): 1–18.

Evola, V., Skubisz, J. and Fernandes, C. 2015. The Role of Eye Gaze and Body Movements in Turn-Taking during a Contemporary Dance Improvisation. *Proceedings from the 3rd European Symposium on Multimodal Communication*. Dublin: MMSYM 2015, 24–31.

Feltham, F., Loke, L., van den Hoven, E., Hannam, J. and Bongers, B. 2014. The Slow Floor: Increasing Creative Agency While Walking on an Interactive Surface. *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction (TEI '14)*. Munich: ACM, 105–12.

Françoise, J., Candau, Y., Alaoui, S. and Schiphorst, T. 2017. Designing for Kinesthetic Awareness: Revealing User Experiences through Second-Person Inquiry. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Denver: CHI '17, 5171–83.

Fuchs, T. and De Jaegher, H. 2009. Enactive Intersubjectivity: Participatory Sense-making and Mutual Incorporation. *Phenomenology and the Cognitive Sciences* **8**: 465–86.

Gill, S. 2015. *Tacit Engagement. Beyond Interaction*. Heidelberg: Springer.

Gurevich, M. and von Muehlen, S. 2001. The Accordiatron: A MIDI Controller for Interactive Music. *Proceedings of the 2001 International Conference on New Interfaces for Musical Expression*. Seattle: NIME, 27–9.

Hall, E. T. 1966. *The Hidden Dimension*. New York: Garden City.

Halm, A. 1978. Musikalische Bildung. In S. Schmalzriedt (ed.) *Von Form und Sinn der Musik*. Wiesbaden: Breitkopf und Härtel, 211–27.

Hanslick, E. 1854. *Vom Musikalisch-Schönen. Ein Beitrag zur Revision der Ästhetik der Tonkunst*. Leipzig: Rudolph Weigel.

Hobeika, L. 2017. Interplay between Multisensory Integration and Social Interaction in Auditory Space: Towards an Integrative Neuroscience Approach of Proxemics. Paris: Université Paris Descartes.

Hove, M. J., Suzuki, K., Uchitomi, H., Orimo, S. and Miyake, Y. 2012. Interactive Rhythmic Auditory Stimulation Reinstates Natural 1/f Timing in Gait of Parkinson's Patients. *PLoS ONE* **7**(3): e32600. https://doi.org/10.1371/journal.pone.0032600

Hunt, A., Wanderley, M. M., Paradis, M. 2002. The Importance of Parameter Mapping in Electronic Instrument Design. *Journal of New Music Research* **32**(4): 429–40.

Jaffe, J. and Feldstein, S. 1970. *Rhythms of Dialogue*. New York: Academic Press.

Keller, P., Novembre, G. and Hove, M. J. 2014. Rhythm in Joint Action: Psychological and Neurophysiological Mechanisms for Real-Time Interpersonal Coordination.

Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences **369**(1658): 20130394. https://doi.org/10.1098/rstb.2013.0394

Kestenberg, J. S. and Sossin, K. M. 1973. *The Role of Movement Patterns in Development*, vol. **1**. New York: Dance Notation Bureau Press.

Kim, J. H. 2013. Shaping and Co-Shaping Forms of Vitality in Music: Beyond Cognitivist and Emotivist Approaches to Musical Expressiveness. *Empirical Musicology Review* **8**(3–4): 162–73.

Kim, J. H., Chang, S. H., Schmidt, L., Otto, J. A., Buch, B., Seifert, U. and Coussement, P. 2010. 'playing_robot': An Investigation of Situated Cognition in the Context of (Artistic) Human-Robot Interaction Design. *Proceedings of IROS 2010 Workshop on Robotics and Musical Expressions*. Taipei: IWRME-2010, 65–72.

Kim, J. H., Reifgerst, A. and Rizzonelli, M. 2019. Musical Social Entrainment. *Music & Science* **2**: 1–17.

Koch, S. C. 2014. Rhythm Is It: Effects of Dynamic Body Feedback on Affect and Attitudes. *Frontiers in Psychology* **5**: Article 537. https://doi.org/10.3389/fpsyg.2014.00537

Kurth, E. 1931. Musikpsychologie. Berlin: M. Hesse.

Livingstone, D. and Miranda, E. 2004. Composition for Ubiquitous Responsive Sound Environments. *Proceedings of the 2004 30th Annual International Computer Music Conference*. Miami: ICMC.

Magnusson, T. 2010. Designing Constraints: Composing and Performing with Digital Musical Systems. *Computer Music Journal* **34**(4): 62–73.

Malloch, S. and Trevarthen, C. 2009. Musicality: Communicating the Vitality and Interests of life. In S. Malloch and C. Trevarthen (eds.) *Communicative Musicality: Exploring the Basis of Human Companionship*. Oxford: Oxford University Press, 1–11.

McArthur, J. A. 2016. *Digital Proxemics*. Bern: Peter Lang.

McGrath, J. E. and Kelly, J. R. 1986. *Time and Human Interaction: Toward a Social Psychology of Time*. New York: Guilford Press.

McHugh, M. L. 2012. Interrater Reliability: The Kappa Statistic. *Biochemia Medica* **22**(3): 276–82.

Mersmann, H. 1925. Zur Phänomenologie der Musik. *Zeitschrift für Aesthetik und allgemeine Kunstwissenschaft* **12**: 372–97.

Pellecchia, M., Beidas, R. S., Mandell, D. S., Cannuscio, C. C., Dunst, C. J. and Stahmer, A. C. 2020. Parent Empowerment and Coaching in Early Intervention: Study Protocol for a Feasibility Study. *Pilot and Feasibility Studies* **6**(22): 1–12.

Pikovsky, A., Rosenblum, M. and Kurths, J. 2001. *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge, MA: Cambridge University Press.

Rabinowitch, T., Cross, I. and Burnard, P. 2012. Long-term Musical Group Interaction Has a Positive Influence on Empathy in Children. *Psychology of Music* **41**(4): 484–98.

Rizzonelli, M., Kim, J. H., Gladow, T. and Mainka, S. 2017. Musical Stimulation with Feedback in Gait Training for Parkinson's Disease. *Psychomusicology: Music, Mind, and Brain* **27**(3): 213–18.

Robson, C. and McCartan, K. 2016. *Real World Research*. Chichester: Wiley.

Robson, D. 2002. Play!: Sound Toys for Non-Musicians. *Computer Music Journal* **26**(3): 50–61.

Rose, D., Delevoye-Turrell, Y., Ott, L., Annett, L. E. and Lovatt, P. J. 2019. Music and Metronomes Differentially Impact Motor Timing in People with and without Parkinson's Disease: Effects of Slow, Medium, and Fast Tempi on Entrainment and Synchronization Performances in Finger Tapping, Toe Tapping, and Stepping on the Spot Tasks. *Parkinson's Disease* **2019**: Article 6530838. https://doi.org/10.1155/2019/6530838

Rosenbaum, D. A. 2010. *Human Motor Control*. Amsterdam and Boston, MA: Elsevier.

Rothfarb, L. 2002. Energetics. In T. Christensen (ed.) *The Cambridge History of Western Music Theory*. Cambridge: Cambridge University Press, 927–55.

Rovan, J. B., Wanderley, M. M., Dubnov, S. and Depalle, P. 1997. Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. *Proceedings of the 1997 AIMI International Workshop. Kansei, The Technology of Emotion*. Genoa: AIMI, 68–73.

Samaritter, R. and Payne, H. 2013. Kinaesthetic Intersubjectivity: A Dance Informed Contribution to Self-other Relatedness and Shared Experience in Non-verbal Psychotherapy with an Example from Autism. *The Arts in Psychotherapy* **40**(1): 143–50.

Searle, J. 1990. Collective Intentions and Actions. In P. R. Cohen, J. Morgan and M. E. Pollack (eds.) *Intentions in Communication*. Cambridge, MA: MIT Press, 401–15.

Searle, J. 2010. *Making the Social World: The Structure of Human Civilization*. Oxford: Oxford University Press.

Seifert, U. and Kim, J. H. 2008. Towards a Conceptual Framework and an Empirical Methodology in Research on Artistic Human-Computer and Human-Robot Interaction. In I. Pavlidis (ed.) *Human-Computer Interaction*. Vienna: In-Tech, 177–94.

Sigrist, R., Rauter, G., Riener, R. and Wolf, P. 2013. Augmented Visual, Auditory, Haptic, and Multimodal Feedback in Motor Learning: A Review. *Psychonomic Bulletin & Review* **20**(1): 21–53.

Stern, D. N. 1985. *The Interpersonal World of the Infant. A View from Psychoanalysis and Developmental Psychology*. London: Karnac Books.

Stern, D. N. 2004. *The Present Moment in Psychotherapy and Everyday Life*. New York: W. W. Norton.

Stern, D. N. 2010. *Forms of Vitality: Exploring Dynamic Experience in Psychology, the Arts, Psychotherapy, and Development*. New York: Oxford University Press.

Stern, D. N., Hofer, L., Haft, W. and Dore, J. 1985. Affect Attunement: The Sharing of Feeling States Between Mother and Infant by Means of Intermodal Fluency. In T. N. Field and N. Fox (eds.) *Social Perception in Infants*. Norwood, NJ: Ablex, 249–68.

Stuart, S. 2012. Enkinaesthesia: The Essential Sensuous Background for Co-agency. In Z. Radman (ed.) *The Background: Knowing Without Thinking*. Palgrave Macmillan, 167–86.

Stupacher, J., Wood, G. and Witte, M. 2017. Synchrony and Sympathy: Social Entrainment with Music Compared to a Metronome. *Psychomusicology: Music, Mind, and Brain* **27**(3): 158–66.

Styns, F., Van Noorden, L., Moelants, D. and Leman, M. 2007. Walking on Music. *Human Movement Science* **26**(5): 769–85.

Tanaka, A. and Knapp, R. B. 2002. Multimodal Interaction in Music Using the Electromyogram and Relative Position Sensing. *Proceedings of the 2nd International Conference on New Interfaces for Musical Expression*. Dublin: NIME02, 43–8.

Thaut, M. H., McIntosh, G. C., Rice, R. R., Miller, R. A., Rathbun, J. and Brault, J. M. 1996. Rhythmic Auditory Stimulation in Gait Training for Parkinson's Disease Patients. *Movement Disorders* **11**(2): 193–200.

Thaut, M. H., Rathbun, J. A. and Miller, R. A. 1997. Music Versus Metronome Timekeeper in a Rhythmic Motor Task. *International Journal of Arts Medicine* **5**(1): 4–12.

Wittmann, M. and Pöppel, E. 1999/2000. Temporal Mechanisms of the Brain as Fundamentals of Communication – With Special Reference to Music Perception and Performance. *Musicae Scientiae* **3**-(1_suppl): 13–28.

Woolhouse, M. H., Tidhar, D. and Cross, I. 2016. Effects on Inter-Personal Memory of Dancing in Time with Others. *Frontiers in Psychology* **7**: Article 167. https://doi.org/10.3389/fpsyg.2016.00167

Zappi, V. and McPherson, A. 2015. The D-Box: How to Rethink a Digital Musical Instrument. *Proceedings of the 21st International Symposium on Electronic Art*. Vancouver: ISEA.

Zuckerkandl, V. 1956. *Sound and Symbol*, trans. W. R. Trask. New York: Pantheon Books.