Gestural Attributions as Semantics in User Interface Sound Design

Kai Tuuri

Department of Computer Science and Information Systems
FI-40014 University of Jyväskylä, Finland
krtuuri@jyu.fi

Abstract. This paper proposes a gesture-based approach to user interface sound design, which utilises projections of body movements in sounds as meaningful attributions. The approach is founded on embodied conceptualisation of human cognition and it is justified through a literature review on the subject of interpersonal action understanding. According to the resulting hypothesis, stereotypical gestural cues, which correlate with, e.g., a certain communicative intention, represent specific non-linguistic meanings. Based on this theoretical framework, a model of a process is also outlined where stereotypical gestural cues are implemented in sound design.

Key words: gestures, user interfaces, sound design, semantics

1 Introduction

Sound-based communication within different kinds of media has a long tradition. Sound design practices for radio-plays from as early as the 1920s have defined the basis for the communicative use of sound effects which is still relevant in today's film and video game sound design [1]. An essential part of the craftsmanship of film sound designers has concerned the creation of sound effects that reflect mental states of the story's characters. The appropriate door knock for film narration, for example, can be *urging*, *gentle* or *angry*, depending on the purpose of that door knock and the feelings and intentions of the person who is knocking. Such focus on *agency* behind sounds is frequently utilised in film sound design. It exploits the perceptual bias towards understanding the human involvement (i.e., *intentionality*) in the sound-causing action [2]. But when it comes to the sound design for human-computer interaction (HCI), such interpersonal focus on interpretation is rarely utilised in a systematic manner. This paper focuses on this agency-orientated perspective.

Due to its history, the HCI field has its roots in information theory [3] and "system-centred" design [4]. Compared to filmmaking tradition, functions of user interface (UI) sounds are more easily conceived in terms of information processing and transmission between machine and user than in terms of interpersonal communication. Such a perspective is well exemplified in the design/research

This is a preprint of an article published in Kopp, S. & Wachsmuth, I. (eds.), Gesture in Embodied Communication and Human-Computer Interaction. Springer, Heidelberg, pp. 257-268, 2010.

paradigm of "earcons" [6], which usually refer to abstract user interface sounds with highly arbitrary meanings. It has adopted a linguistically orientated view of semantics which usually sees the semantic content as symbolic units of information essentially separable from its form of expression. The risk of such an approach is that design can become detached from meaningful experiences that get coupled with sounds in the interaction. It should be obvious that the role of sound cannot be as a mere carrier of symbolic information. Indeed, reflecting the ongoing shift towards user-centred design [5], contemporary trends in HCI sound research have preferred to talk about *sonic interaction design* thus emphasising the coupling of sound and its meanings with interaction [7].

The word *gesture* is used here to represent any bodily act that – observable in interaction – operates as a vehicle for interpersonal communication. Such a perspective is not restricted to hand movements, but takes all non-verbal forms of body-related communication into account. In social interaction, we express our mental states with bodily actions which can be either directly perceivable (like in hand/facial gestures or in vocal prosody), and/or indirectly perceivable as reflections of body movements (e.g., in sounds of objects which are acted on). Communication with gestures is primordial [8] and often unconscious for us. The basis for gestural communication – the physical constitution of the human body and our ways to schematise it – is universal. Gestural communication has also been suggested to have a strong phylogenetic background which precedes verbal communication [8]. As Marc Leman has put it: "Gestures form the basis of mutual adaptive behavioral resonances that create shared attention and are responsible for the feeling of being unified with other people" [2].

By suggesting the utilisation of gestural attributions (projections of motor-activity/body movement) as semantics in UI sound design, this paper emphasises bodily mediated action understanding and the role of action-relevant gestural cues in sound as constituents of meaning-creation based on a kinaesthesic foundation. In the scope of interaction design, such tacit "sensibility for movement" also accounts for a "sensibility for responses to movement" [9]. We thus stress the close engagement of interaction and meaning-creation already acknowledged within, for example, the ecological view [10] of perception. Gestural attributions, unlike linguistic ones, are not detached from the direct sensory-motor basis of social interaction. According to the embodied approach to human cognition [11], the human mind is coupled with our environment. That coupling with the environment has emerged in the course of the experiential history of using our bodies for interacting with it. Understanding is thus inseparable from the embodied experiences of the physical world including – most relevantly to this study – interactions with other people.

The aims of this paper are to promote an idea that gestural projections of body movement attributed to sounds could be used as semantics for UI sound design, and to formulate this idea into a justifiable and testable hypothesis. In addition, the aim is to outline a model of the process in which gestural cues are implemented in sound design.

2 Embodied Basis of Understanding Gestures

In the phylogenesis of social mammals, such as humans and non-human primates, it has proved to be beneficial – maybe even essential for survival – to understand the actions of others [12]. Surely, without such an ability, the social and cultural development of humans would have been impossible. In this section we discuss the embodied basis of this attunement to interpersonal relations, which involves shared body-related constituents. First, the neurological foundation of *interpersonal action understanding* is reviewed, which is then applied to the concept of body-schema and empathetic involvement in perception. Lastly, interpersonal action understanding is viewed in terms of the Brunswikian lens model.

2.1 The Human Mirror-Neuron Mechanism

Mirror neurons are a particular class of premotor neurons that discharge both when one performs a specific goal-reaching action and when one observes other individuals executing similar actions [12]. They were originally discovered in the monkey premotor cortex, but there is also evidence for the existence of a similar mirror system in humans [12].

Nowadays there exist two parallel and equally plausible hypotheses about the functional role of mirror neurons. Firstly, they mediate bodily imitation and secondly, they are related to action understanding [12]. The motor representation encoded in mirror neurons thus reflects the understanding of observed action – not object presentation. Via motor representation, mirror neurons transform sensory information into knowledge that agrees with the motor repertoire/skills of the observer [12]. In other words, the observer understands the performed action as she could perform it herself. Action understanding thus involves embodied "resonances" (or embodied simulation [13]) of the observed action.

Experiments have shown that even fragmentary clues about action presented to the observer can trigger the specific response (motor representation) in mirror neurons [12]. Therefore, the audio-visual features of the observed actions seem to be fundamental only to the point where they allow action understanding. For example, the mere sound of action seems to result in a response that matches the responses for the same action observed or executed [14]. The encoding of action in the mirror system thus seems to be highly multimodal in nature. These above aspects underline the possible role of mirror neurons in contributing schematic gestalt processes, which transforms sensory information (like hearing somebody laughing) into preconceptual structures meaningful to the observer (understanding laughing by means of mirrored motor representation of it). It is also suggested that such action understanding operates as an enabling mechanism for empathy [13].

2.2 Interpersonal Body-Schematic Transfer

The concept of *body-schema* refers to a tacit understanding of one's own body in-the-world. As suggested in *Phenomenology of Perception* by Merleau-Ponty [15], humans possess such specific schemata of our body in relation to embodied space, i.e., space in the environmental setting of our habitual actions. It is reasonable to assume that such body-schemata (or kinaesthetic image schemata [16]) are based on recurrent sensory-motor experiences of bodily interactions with the world.

According to Merleau-Ponty [15], body-schema has a crucial function in the perception of other individuals as human-beings. That function is related to *body-schematic transfer* where the movements of other individuals are perceived as the movements that the observer could imagine executing by her own action repertoire. Therefore the perception of body movements is based on the perceiver's embodied knowledge of body-schema. This theoretical idea is very much in line with the already discussed function of mirror neurons in action understanding. It is thus plausible that the mirror system is a part of the realisation of body-schematic transfer [17].

Jan Almäng [17] has proposed that, at its basic level, body-schematic transfer has at least four characteristic features, which are:

1. The perceiver observes the other as having a body-schema.
2. The perceiver can perceive the action by means of body-schema even when she is unable, e.g. due to its complexity, to perform it herself. Thus, it is sufficient that her body-schema can "read" the movement.
3. Physical similarity between the perceived and perceiver is not required for an apprehension of the movements by body-schematic transfer. Thus, it is sufficient if there appear to be kinematic similarities between observed movement and body-schematic knowledge of how to produce such movement.[1]
4. Body-schematic transfer of movements by itself does not imply that the intentionality of the other person is communicated to the perceiver. This is because understanding the mental states of someone, on the basis of physical movement, requires contextual awareness.

Because human actions arise from intentions, emotions and other affective determinants which are linked to the context, we must situate movements in the interaction where it takes place. This broader aspect of body-schematic transfer is discussed in the following section.

2.3 Empathetic Involvement in Perception

We normally perceive other people as engaged in situations that provide meaningful references to their actions. Hence, the crucial element in the Merleau-Pontyian view of understanding others (by means of body-schematic transfer) lies in the ability to automatically re-center our perception of the situation to the perspective of the other. Therefore, together with tacit perception of a person's body movements, the observer re-centers her own primordial perception

[1] This feature is evident in watching animated cartoons, where even objects can indicate such anthropomorphism in movements that they can be perceived by means of body-schematic transfer.

of the surrounding environment – including its action affordances – to become a perception *for* that other person [17, 13]. In this way we can understand motivations and other reflections of intentionality in actions, and we are also able to anticipate possible actions of another person – as if they were actions of our own. Such an attunement to another individual, not only to body movements but also to her intentionality and action affordances, seems to be such a natural part of our interpersonal awareness that it requires no conscious reasoning.

To sum up the discussion on the embodied understanding of movements, we can assume that it is based on two parallel aspects in perception. The first aspect is related to the mimetic involvement in the mirror neuron system, by which perceived movement is understood in terms of body-schema and kinaesthesia. Such corporeal resonances can range from simple synchronisations to more specific motor mimetic attuning. The second aspect is related to empathetic involvement, in which the body-schematic ”resonances” (of the first aspect) are associated with the perspective of the other individual engaged in interaction. At the lower level, this refers to mere action-based involvement with the other's movements and thus primordial apprehension of *corporeal intentionality* (i.e., motor intentionality) [2], whereas ”genuine” empathy usually refers to more participative, emotional and inferential involvement with the perspective of the other. A similar distinction, between the degrees of motor-system involvement and the degrees of empathetic involvement, has also been suggested in the theory of bodily mediated experiences in music [2].

2.4 Encoding and Decoding of Gestural Cues

We finally take a look at the Brunswikian lens model scheme initially proposed as a framework for understanding how prosodic cues are encoded in a vocal expression of emotion [18]. Based on the original lens model [19], it describes the processes of various *distal cues* being situationally determined in articulation, indicating affective states of a person, and how these acoustically transmitted patterns (as *proximal cues*) play a role in the attribution of an affective state in perception. The lens model simultaneously considers *encoding* (i.e., contextual determination of cues) and *decoding* (i.e., contextual interpretation of cues). It therefore gives a neat overview of communication, where body is acknowledged as a mediator. The model, adapted for gestural communication, is illustrated in Figure 1.

In the encoding of cues, Scherer has emphasised the central role of *push* and *pull effects* [18]. Therefore intentional and affective states of a subject are situationally characterised in gestural articulation by interaction between 1) psychobiological processes, intrinsically related to mental states, that provide a natural influence on body movements (push effects) and 2) interactional processes, which involve voluntary control over body movements and are related to external conditions (pull effects). Conditions of interaction thus often requires a certain strategic display (or hiding) of intentions and other mental states.

The dominance of push effects is most evident in so-called *affect bursts*, which are mostly a result of physiological arousal. Push effects also have significant role

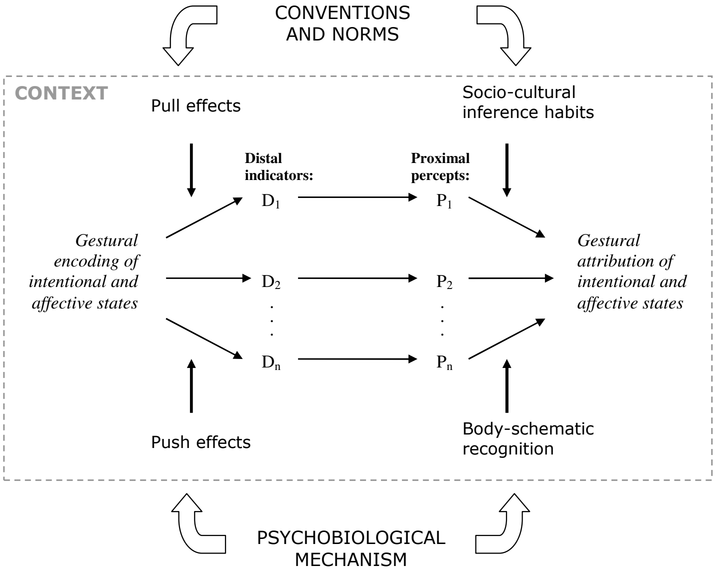

Fig. 1. The Brunswikian lens model with push and pull effects [18].

in spontaneous expressions of, for example, pain or joy, but they also influence even the most premeditated acts of communication. [18] At the other end of the continuum, the dominance of pull effects is evident in expressions of an intended purpose. John R. Searle has suggested three *conditions of satisfaction* that fuel speech acts [20], but they should also apply to any acts which bear the motivation to be understood in interaction. As determinants of gestural articulation they are 1) *articulation intention*; an intention to appropriately produce a certain kind of gesture, 2) *meaning intention*; an intention to mean something with the gesture, and 3) *communicative intention*; an intention to be understood in a certain way – referring to the type of communicative intent such as *asking*. Communicative intention essentially imposes its condition of satisfaction on the conditions of satisfaction of both meaning something and producing a gesture accordingly.

According to the lens model, distal cues cannot be directly perceived. But on the basis of body-schematic transfer, the proximal cues can be perceived in terms of distal gestural movements. Proximal cues are probabilistic, partly redundant and contribute to action understanding in an additive fashion [21]. Gestural action understanding, in turn, provides cues for attributing intentionality [2]. Gestural attributions are thus contextually constructed in interaction between primordial, ecologically developed gestalts and inference based on social habits or cultural norms. These two processes can be seen as perceptual counterparts to push and pull effects, in which proximal percepts of gestural cues resonate with both body-schematic and conventional backgrounds.

3 Gestural Attributions in Sound Design

This section concentrates the discussion around a gesture-based approach to UI sound design – the principle of using sounds as the means of projecting gesturally attributable cues of movement. The basic assumption in such an approach is that human capacity for interpersonal awareness allows body movements of another individual to be understood in terms of the body movements of the perceiver. This seems to apply also in situations where the presented movement is partial or, due to its kinematic similarities with body-schematic reference, merely implies bodily movement. As already discussed earlier in the paper, action understanding is not dependent on which sensory modality is utilised in providing clues of action. Therefore the premise is that sounds can be used for presentation of action-related features.

The lens model perspective (see 2.4) demonstrates how perceptual action understanding is an emerging resonance of the mixture of several parallel cues encoded by the same gesture. In the context of HCI sound design, the strength of the lens model is that it outlines successful communication by means of probabilistic, multiple and redundant cues that allow discarding non-relevant cues for the task. Due to technical or aesthetic reasons, the designer must often conform to a selected set of acoustic cues in attributing intended characteristics to UI-sounds. Limitations should not undermine the utilisation of gestures, as there should be plenty of suitable cue combinations which compensate for the discarded ones. Indeed, it has been found that even simple acoustic cues can communicate the emotions of a musical performer or a speaker (with a general lack of cue interactions) [21]. As interaction is multimodal, the action understanding is ultimately based on the contextual whole, in which the sound instance and its cues become perceptually fused with the other aspects of interaction [22].

3.1 Gestural Articulations Projected in Sounds

From the perspective of the observer/user, gestural projections of sound refer to kinaesthetic imagery of body movement, which arise during the listening experience. The contextual creation of such gestural imagery is based on body-schematic resonances of *motor-mimetic* involvement [23] in listening. It is easiest to assume that gestural imagery is perceived in sounds that are in some way caused by bodily excitation, hence implying sound-producing gestures. But regardless of the type of sound production, motor-mimetic involvement applies to the perception of music, or *any* sound, as long as it is able to imply physical movement [23, 2]. Thus, gestural imagery can even be attributed to abstract and artificially produced sounds.

From the sound designer's perspective, communication with gestural projections means defining a contextually appropriate gesture for the communicative purpose, and then articulating and implementing it in design. Gestural articulation can be a vocal act, musical expression or any physical action that itself produces, or allows its features to be transformed into, acoustic resonances. As discussed in 2.4, the articulation is bound up with the situation. When a gesture is articulated spontaneously, while being immersed in interaction, articulation is not a subjective interpretation so much as an *experience as articulated*. A sound, caused or modulated by gestural articulation consequently conveys acoustic cues of corporeal intentionality involved *in* the physical articulation.

The basic principle of using gestures as part of sound production in HCI is not new, although the idea has usually been utilised in producing immediate environmental audio feedback on the basis of the user's gestures (e.g. [24]) – not in exploiting gestures as interpersonal communication. It can be argued that traditional film sound design practices have long acknowledged the importance of gestural articulation in creating sound effects. One prominent example of this is a tradition called *foley art* [1]. Despite all the sophisticated audio technology available today, foley art still favours manual ways of producing sounds (in real-time). As noted above, gestural communication also refers to acting on material objects. This is exactly the case with foley art, in which the aim is to express through the sounds of material objects and provide "added value" to the narrative whole. We see that direct bodily involvement during sound creation – often performed concurrently with the related visual narration – enables the intentionality of a performer to be communicated via gestural projection. As illustrated in the introductory section of this paper, even simple sounds like door knocks can have much variety both in their gesturally determined qualities and in how these qualities can affect the contextual interpretations (see also [25]).

3.2 Utilising Stereotypical Gestural Cues in Sound Design

When treating gestural attributions as semantics, the sound designer can approach them from at least two directions: She can use the gesture as a starting point (thus emphasising distal cues), also accounting for the situated articulation and motivation of the gesture. Or she can focus on gesture-specific acoustic characteristics (thus emphasising proximal cues), but only if she has sufficient knowledge about the acoustic correlates of gesture-related understanding. Thus we regard it as easier for the designer to start with communicationally appropriate gestural articulation as the means to acquire and study gestural cues, which in turn can be utilised in sound production. In this way, semantics is always considered as being closely linked to the context.

In order to use gestural semantics systematically, one needs a way to categorise different types of meanings, i.e., gestural patterns that become contextually meaningful. In order to take advantage of the non-linguistic characteristics of gestures, we are especially interested in discovering *stereotypical* gestural cues. They reflect embodied meanings of a specific type of *recurrently experienced* gesture and resonate with primordial gestalts for interpersonal understanding and communication. Stereotypical cues are thus a type of gesturally perceived cues that should communicate a specific meaning robustly, without being too strongly dependent on cultural constraints. They differ from "weak cues" (like indications of direction or force), which are extremely dependent on context. They also differ from "coded cues", which communicate robustly, but are only meaningful because of coding or convention.

But how can the sound designer find gestures that convey such stereotypical meanings? We suggest using the *communicative intention* of gesture as a category of semantics. The assumption is that context-situated articulation – with specific intention to communicate – results in stereotypical physical cues of that intention as an outcome of push and pull effects. There is evidence, for example, that prosodically realised acoustic patterns specific to communicative intention exist [26, 27]. Arguably, in infant-directed speech, such intention-specific prosody (e.g., for alerting or prohibiting) functions as the first regular semantic correspondence to the infant, clearly preceding any linguistically related functions of prosody [26]. Communicative intention of gestural expression thus appears to serve a fundamental and important prelinguistic function.

Prosodic patterns represent a subset of gestural patterns as they are caused by "phonetic gestures"; motor movements of the phonatory apparatus, vocal tract and respiration. Indeed, there are two reasons why prosody of vocal acts promises to be a very important source for gestural cues being utilised in sound design. Firstly, the evolution of the vocal apparatus is related to human communication [26]. Secondly, in prosody, gestural cues are directly realised as acoustic cues which are familiar and ecologically valid to us. Of course, stereotypical cues – specific in communicating different emotions or interpersonal attitudes – exist in other kinds of gestural expressions as well [28]. But in order to be encoded acoustically, non-vocal gestures need to be sonified. In everyday interactions, such sonification is a natural outcome of material resonances of motor excitation when objects and materials are acted on. However, by using physical models (e.g. [24]), the sounds of various material interactions can also be synthesised on the basis of, e.g., kinematic parameters.

Figure 2 specifies the general phases of *modelling*, *performing*, *utilising* and *evaluating* in the process of implementing gestural cues in sound design. In the first phase, the need/purpose of an UI sound element is acknowledged on the basis of the designer's model of application-user interaction. Hence the communicative functions for UI sounds can be determined. The modelling of appropriate gestural action requires mental exploration of interaction (see the discussion about action models in [22]). If the designer puts herself into the dialogue between user and machine, she is able to conceive her role as a person who is communicating with the user. The designer can thus imagine participating in the interaction, which in reality occurs via the mediation of the machine in use. From that perspective, she can mentally explore the patterns of contextual application use and – whenever sonic feedback is required – discover gestural patterns that would feel contextually appropriate for the communicative need. The communicative intention of the modelled gesture thus conforms to the communicative function of the propositional UI sound.

In the performing phase, the specified gesture is articulated. In order to achieve the spontaneity in articulation, the gesture should be performed while being immersed in interaction. To enable such immersion, a suitable scenario can be used providing the situational flow of interaction. This can be done, for example, in terms of metaphorical person-to-person interaction (see example

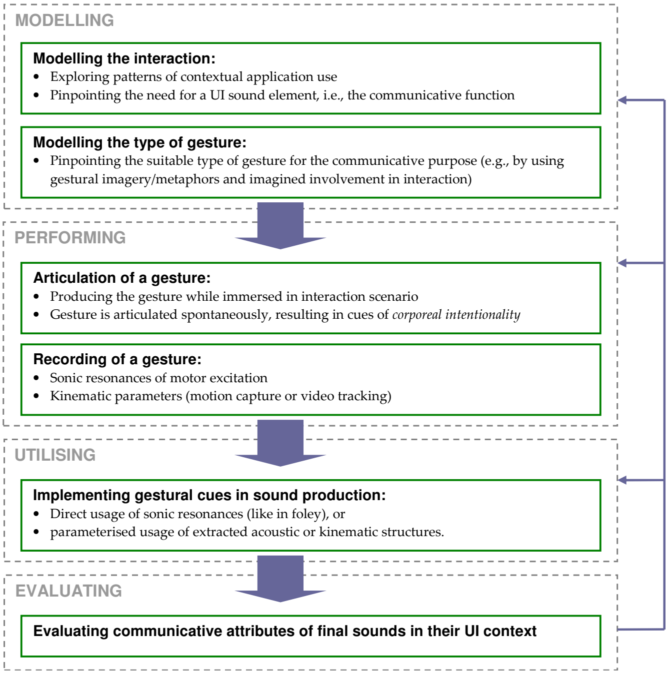

Fig. 2. The process of implementing gestural cues in sound design.

cases in [27, 22]). The articulation of a gesture needs to be recorded, either as a sound recording (of sonic resonances of action) or as a recording of selected kinematic parameters (by motion capture).

The utilisation of recorded gesture in sound production can be based on either ”foleyish” direct usage of recorded sound, or – in a more analytical manner – parameterised usage of gestural cues. Parameters can be acquired, for example, by extracting selected acoustic features from sound recording (see example cases in [27, 22]), or by utilising recorded kinematic parameters of gestural articulation. Parameterised gestural cues can be implemented in sounds, for example, as parameters for sound synthesis or sound manipulation. Alternatively they can be used indirectly as structural ideas in sound design and production. The last phase in the process is evaluation, where the communicative attributes of the final UI sounds are contextually tested. When required, the sound designer can iterate and re-evaluate the process starting from any of the previous phases.

4 Conclusions

We can now conclude this paper in the form of a hypothesis about using gestural projections of body movement as semantics in UI sound design. The resulting hypothesis includes the following assumptions:

- Sound design can be founded on theory which emphasises embodied cognition and interpersonal action understanding.
- By exploiting the primordial capacity for interpersonal action understanding in humans, sound design can utilise stereotypical, gesturally realised cues of social interaction, which represent non-linguistic categories of meaning.
- On the basis of the theoretical framework presented, sound design can be outlined as a process which is explicated into distinct design phases where action-relevant cues are determined in terms of interaction and gestural articulation.
- The approach presented results in ecologically valid semantics which should be communicated robustly and require less learning than semantic attributions of linguistically orientated design schemes. The ongoing research, based on the theoretical framework presented, has supported this assumption [29].

The gesture-based perspective on sound design provides an important focus on performative aspects of sound design (i.e., direct involvement with the sound creation), and bodily engagement with sonic communication (both in sound design and contextual perception). These aspects merit more explicit focus within sound design research, although they are most likely tacitly acknowledged by many professional sound designers. A linguistically orientated paradigm often considers semantics as "absolute". Such a perspective easily dismisses the communicational potential that even a simple feedback sound can have when it is designed as physical activity – and for physical activity.

Acknowledgments. This work is funded by Finnish Funding Agency for Technology and Innovation, and the following partners: GE Healthcare Finland Ltd., Suunto Ltd., Sandvik Mining and Construction Ltd. and Bronto Skylift Ltd.

References

1. Mott, R. L.: Sound effects: Radio, TV and Film. Boston, MA: Focal Press (1990)
2. Leman, M.: Embodied Music Cognition and Mediation Technology. Cambridge, MA: MIT Press (2008)
3. Shannon, C. E. and Weaver, W.: The Mathematical Theory of Communication. Urbana, IL: The University of Illinois Press (1949)
4. Czaja, S. J.: Systems design and evaluation. In Salvendy, G. (ed.) Handbook of human factors and ergonomics (2nd ed.). New York, NY: Wiley, 17–40 (1997)
5. Bannon, L. J.: A human-centred perspective on interaction design. In Pirhonen, A., Isomäki, H., Roast, C. and Saariluoma, P. (eds.) Future interaction design. Berlin/Heidelberg: Springer-Verlag, 31–51 (2005)
6. Brewster, S., Wright, P. and Edwards, A.: A detailed investigation into the effectiveness of earcons. In Kramer, G. (ed.) Auditory display. Reading, MA: Addison-Wesley, 471–498 (1994)
7. Rocchesso, D., Serafin, S., Behrendt, F., Bernardini, N., Bresin, R., Eckel, G., Franinovic, K., Hermann, T., Pauletto, S., Susini, P and Visell, Y.: Sonic interaction design: sound, information and experience. In CHI '08 extended abstracts on Human factors in computing systems. New York, NY: ACM press (2008)

8. Rizzolatti, G. and Arbib, M. A.: Language within our grasp. Trends in Neurosciences, 21, 188–194 (1998)
9. Svanaes, D.: Kinaesthetic thinking. Computers in Human Behavior, Vol. 13, No. 4, 443–463 (1997)
10. Gibson, J. J.: The ecological approach to visual perception. Boston, MA: Houghton Mifflin (1979)
11. Varela, F., Thompson, E. and Rosch, E.: The Embodied Mind. Cambridge, MA: MIT Press (1991)
12. Rizzolatti, G. and Craighero, L.: The mirror-neuron system. Annual Review of Neuroscience 27, 169–192 (2004)
13. Gallese, V., Eagle, M. N. and Migone, P.: Intentional Attunement: Mirror Neurons and the Neural Underpinnings of Interpersonal Relations. J Am Psychoanal Assoc, Vol. 55, No. 1, 131–175 (2007)
14. Kohler E., Keysers C., Umiltá MA., Fogassi L., Gallese V. and Rizzolatti G.: Hearing sounds, understanding actions: action representation in mirror neurons. Science 297, 846–848 (2002)
15. Merleau-Ponty, M.: Phenomenology of Perception. London: Routledge (1945/1996)
16. Johnson, M.: The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason. Chicago, IL: University of Chicago (1987)
17. Almäng, J.: Intentionality and intersubjectivity. Ph.D. thesis, Goteborg, Sweden: Goteborg Universitet (2007)
18. Scherer, K. R. and Bänziger, T.: Emotional expression in prosody: a review and an agenda for future research. In proc. of SP-2004, Nara, Japan, 359–366 (2004)
19. Brunswik, E.: Perception and the representative design of psychological experiments (2nd ed.). Berkeley, CA: University of California Press (1956)
20. Searle, J. R.: Mind, language and society. New York, NY: Basic Books (1998)
21. Juslin, P. N. and Laukka, P.: Communication of emotions in vocal expression and music performance: Different channels, same code? Psych. Bull., vol. 129, no. 5, 770–814 (2003)
22. Tuuri, K., Pirhonen, A. and Välkkynen, P.: Bodily Engagement in Multimodal Interaction: A Basis for a New Design Paradigm? In Kurkovsky, S. (ed.) Multimodality in Mobile Computing and Mobile Devices. Hershey, PA: IGI Global (2009)
23. Godøy, R. I.: Gestural-sonorous objects: embodied extensions of Schaeffer's conceptual apparatus. Organised Sound, 11(02), 149–157 (2006)
24. Avanzini, F. Serafin S. and Rocchesso, D.: Friction sounds for sensory substitution. In proc. of ICAD 2004, Sydney, Australia (2004)
25. Vitale R. and Bresin, R.: Emotional Cues in Knocking Sounds. In proc. of ICMPC 10 (abstract only), Sapporo, Japan, p. 276 (2008)
26. Fernald, A.: Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In Barkow, J., Cosmides, L. and Tooby J. (eds.) The Adapted Mind: Evolutionary Psychology and the Generation of Culture. Oxford University Press, 391–428 (1992)
27. Tuuri, K. and Eerola, T.: Could function-specific prosodic cues be used as a basis for non-speech user interface sound design? In proc. of ICAD 2008, Paris, France (2008)
28. Feyereisen, P. and de Lannoy, J.-D.: Gestures and speech: Psychological investigations. New York, NY: Cambridge University Press (1991)
29. Tuuri, K., Pirhonen, A. and Eerola, T.: Design and Evaluation of Prosody Based Non-Speech Audio Feedback for Physical Training Application. (journal submission)