Intelligence Technologies as a Means of Enhancing Spatial Experience

Charalampos Rizopoulos[1], Dimitrios Charitos[2]
Laboratory of New Technologies in Communication, Education and Mass Media
Department of Communication and Media Studies,
National and Kapodistrian University of Athens
[1]harisrizopoulos@yahoo.gr, [2] vedesign@otenet.gr

Following the evolution of human-computer interaction to date, intelligent user interfaces (IUIs) seem to be one of the most important paradigms for future research. "Intelligent" or quasi-intelligent behaviour may be applied to both virtual (intelligent virtual environments) and real space (Ambient Intelligence). This paper focuses on the augmentation of physical space as a characteristic of the latter case. More specifically, this paper briefly examines Ambient Intelligence and the concepts on which it is based, documents the importance of Ambient Intelligence (AmI) technologies for the formulation of the user's spatial experience and attempts to outline some important theoretical approaches to human behaviour and communication that need to be considered during the design, implementation, and evaluation of Ambient Intelligence systems. Ultimately, this paper aims to outline the influence of said technologies on the users' activity within the environment and their environmental experience in general.

Keywords: *Ambient Intelligence; environmental design; activity theory; adaptation.*

Introduction

This paper documents the current state of a research project the goal of which is the development of a new theoretical model that will support the design and implementation of Ambient Intelligence (AmI) environments from a communicational perspective. More specifically, the paper reviews relevant literature and attempts to formulate hypotheses on the characteristics of this model. These hypotheses will be subsequently tested during the course of the project. Firstly, Ambient Intelligence and the concepts on which it is based are being investigated in order to identify the relevance of these technologies to the transformation of the user's spatial experience. Secondly, certain important theoretical approaches to human behaviour and communication are also considered in an attempt to identify their relevance for informing the design, implementation, and evaluation of Ambient Intelligence systems.

Ambient Intelligence

Ambient Intelligence (AmI) is a relatively new field of ICT research that has emerged out of the convergence of three other technologies: ubiquitous (also called "tangible" or "pervasive") computing, multimodal communication, and intelligent user interfaces. Its ultimate goal is to support and augment human-computer interaction by positioning this activity within an everyday spatial context. In contrast to Virtual Reality (VR), the user is not placed inside a synthetic environment; rather, computers and their input/output devices are placed in the real environment the user inhabits. The user does not adapt to the computer; instead, the computer adapts to the user through the support of "intelligent" techniques. More specifically, AmI offers the possibility of embedding networked intelligent artefacts in real environments, thus inserting a layer of information with which users interact in the context of an enhanced environmental experience. This experience is no longer only determined by static spatial elements but also consists of intelligent artefacts, displaying multisensory content, the course of which may be dynamically altered as a result of user interaction. Therefore, the user's experience of an AmI system is radically different to the experience of using an ordinary desktop computer.

Some authors (Weiser, 1991; Riva, 2005) consider AmI as the exact opposite of VR. This assumption is largely based on the fact that AmI and VR, as mentioned above, pertain to very different kinds of a spatial experience. However, these two technologies may also share one of the most important elements of future trends in interface research and that is their adaptive nature. Thus, for the purpose of this paper, AmI and VR are considered the two sides of the same coin. Consequently, although this paper concentrates mainly on AmI, some of the theories described herein may also apply to intelligent virtual environments. This view is outlined in Figure 1.

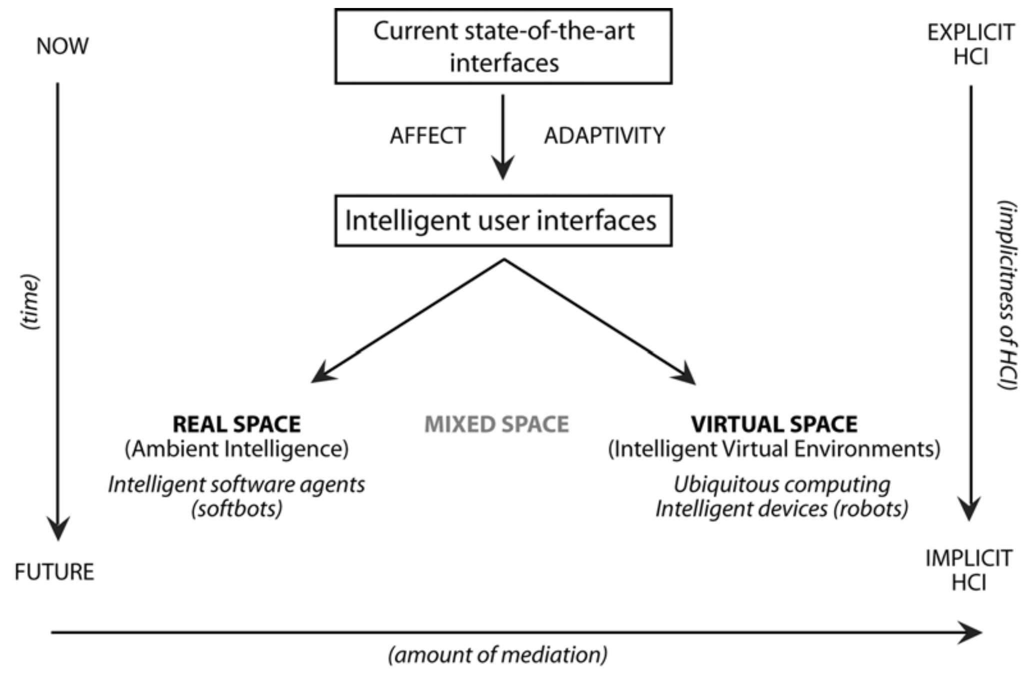

Figure 1
Future trends in human-computer interface research

Relevance of certain future trends in interface research

Intelligent systems receive input via sensors, process this information, and generate appropriate output based on their capabilities, the task to be accomplished, environmental circumstances, etc. The preceding discussion of AmI reveals that future research on interfaces revolves around three axes: ubiquity, multimodality, and adaptation.

Ubiquity

The concept of "*Ubiquitous computing*" was coined by Mark Weiser (1991) and refers to the implementation of numerous, possibly wireless, networked devices, scattered throughout the physical environment. It takes into account the fact that humans are physical beings "unavoidably enmeshed in a world of physical facts" (Dourish, 2001, p.99) and thus familiar with physical, real world artefacts.

An essential aspect of ubiquitous computing is the "disappearance" of the devices used, so that the user does not perceive them for what they are. This disappearance may occur on two levels. The first level is that of physical disappearance and may be achieved through the use of sufficiently small devices or other means aiming towards "hiding" the devices from plain sight. Furthermore, ubiquitous computing entails the "decentralisation" of interaction: the user interacts with the system using many different devices across the environment. Thus, there is no "central" point of interaction, as in the case of more conventional systems like a desktop computer (Dourish, 2001). The second level is mental disappearance, whereby the user perceives computing devices as another object with enhanced features[1]. This characteristic has been called "transparency" and is described as the invisibility of technology due to its integration into the general ecology of the environment (Alcapiz and Rey, 2005, p.4). The user is presented with an environment which offers a wider selection of possible actions while retaining the familiar structure he is accustomed to. Therefore, ubiquitous computing affords the embedment of intelligent artefacts in real environments in such a way that the interface is spatialised.

Multimodality

This characteristic refers to the ability of the system to receive multisensory/multimodal input and offer similar output. Multimodal input and output may include facial expression recognition, gaze direction detection, natural language processing, haptic feedback, brain-actuated interaction etc.

Multimodality is regarded as a very efficient way of interacting with a computer. The human body and its senses have been regarded as the "ultimate computer interface" (Krueger, 1993, p.36), and Negroponte (1996) considers multimodality as a key characteristic of a usable system. Humans utilise various modalities during direct human-human communication, such as gestures, body language, vocal intonation etc. The implementation of such modalities for the purpose of human-computer interaction results in less cognitively demanding interfaces (Russell et al., 2005). Furthermore, the inclusion in the interaction process of other senses besides vision, which is currently dominant, may urge users to adapt faster to a system (Kuivakari and Kangas, 2005) or to disregard certain types of technical deficiency (Negroponte, 1996).

Experiencing space does not only involve the perception of visual sensory input but of auditory, olfactory, thermal and tactile input, along with the sense of proprioception - as defined in Gibson (1986, p.111) - all of which play their part in the establishment of a sense of space. This approach is also in agreement with Hall's conception of space (1966, pp.41-63). Therefore, it could be suggested that the development and use of multimodal interfaces may result in spatial interfaces affording a more complete spatial experience.

[1] an interactive table is primarily a table with extra capabilities, not a computer that has merely assumed the form of a table (Streitz and Nixon, 2005).

Adaptation

The third and probably most important axis around which research on future interfaces revolves is that of adaptation. An adaptive system uses appropriate AI methods, such as intelligent agents, adapts to the user's characteristics and environmental conditions and constantly strives to improve itself by utilising appropriate AI techniques. As Dourish (2001) remarks, context has been the focus of ubiquitous computing all along, as a means of action disambiguation. Advances in adaptive technology will enable the design and production of truly "personal" computers and personalised interfaces. The incorporation of Affective Computing in AmI will help address this issue more effectively.

Affective Computing

Since the stated goal of AmI is the support of users' activities and interactions (Riva, 2005, p.19), it only seems natural to take the emotions of users into account. Compared to "mainstream" Artificial Intelligence research, this is a novel approach. Emotions had been considered an impediment to reason and were thus left out of the scope of AI. However, it has been shown (Picard, 1997; Goleman, 2004) that the regulatory role of emotions is very important to what is commonly considered "intelligent" or rational behaviour. Therefore, the fact that AI had not paid sufficient attention to emotions seems paradoxical.

In an attempt to address the role of emotions in computing, Picard (1997, p.3) coined the term "Affective Computing" to describe "*computing that relates to, arises from or deliberately influences emotions*". The objective of Affective Computing is to improve the quality of interaction between the user and the computer. As such, it has been argued (Alcañiz and Rey, 2005) that Affective Computing should be an essential component of AmI.

The expression of human emotions may become accessible to computers through the use of appropriate sensors[2]. A detailed description of human emotion signal capture and processing is, obviously, beyond the scope of this paper, but some indicative means that might be utilised by an affective system include: facial expression extraction, gesture recognition, vocal intonation analysis, and certain physiological signals, such as blood pressure, skin conductivity, pulse, respiration rate, perspiration, pupil size etc.

It should be noted that Affective Computing merely aims to make the computer responsive to the emotions of users and not to make computers "have" emotions as humans do. Such a thing may be impossible or undesirable, whereas allowing the computer to imitate emotionality is a more feasible and less dangerous goal (Picard, 1997).

Towards an implicit interaction process

The combination of the aforementioned tendencies and technological advances points to the replacement of the dominant, dialogue-like human-computer interaction mode. The utilisation of the methods described earlier will allow for the adoption of a different, more implicit form of human computer interaction, which Schmidt (2005, p.164) appropriately names "Implicit Human-Computer Interaction" (iHCI) and describes as the interaction of a human with the environment and artefacts found in it towards the accomplishment of a goal. The user offers implicit input and receives implicit output. Implicit input refers to the actions and behaviours of the user, which are not considered primarily as initiating interaction, but are perceived by the system as such. Implicit output, similarly, refers to output, which occurs as a result of the reception and processing of implicit input. Implicit output is seamlessly integrated with the environment and supports the user's task.

In the context of iHCI, the interaction process is transformed from unidirectional (user-initiated) to bidirectional (mixed-initiative): the system gains the ability to detect subtle communicational cues inherent in the behaviour of a human through the use of appropriate devices that allow for the capture of multimodal input data. By processing these data,

2 There is an obvious link between multimodality and the computer's affective awareness.

the system reaches some conclusions about the user's state and the task to be accomplished. Eventually, the system may subtly act on the environment, aiming at increasing the possibility of the user successfully completing the task.

Investigating user experience and communication in AmI environments

The advent of IUIs is expected to substantially alter the relationship between user and computer. Although humans play a key role in human-computer interaction, the majority of research on the future of interfaces still concentrates on technical aspects, even though the goal is the computer's adaptation to the user. A more human-centred (anthropocentric) approach would be to learn from theories that focus on the study of human-human communication and human behaviour within a given setting in order to strive for a better adaptation of computer systems to the user's needs and preferences.

From a communicational perspective, AmI is expected to alter a person's experience of reality by interposing an additional layer of mediation between the user and the environment. This layer may clearly have an impact on users' conception of the computer and on user's behaviour within such an enhanced environment. This section will focus on theories that have attempted to describe the relationship between the user and the context[3] where interaction takes place in, in order to develop a better understanding of AmI technologies at a communicational level and ultimately to inform the design of intelligent environments.

Paradigm shift: the computer as a potential interlocutor

The cognitive approach is the earliest paradigm that describes the relationship between human and computer. According to this approach, the human brain is seen as a specific type of an information-processing unit consisting of a sensory input subsystem, a central information processing subsystem, and a motor output subsystem (Kaptelinin, 1996). Memory provides a comparison with the representations of past experiences, thus aiding the processing. Furthermore, an action may take place in the cognitive conscious or the cognitive unconscious, depending on whether it can be performed automatically, without conscious effort (Raskin, 2000).

Human-computer interaction is seen as a constant loop involving two systems, the human and the computer. One's output is used as the other's input, and so on. While this approach may have the advantage of simplicity, it fails to take into account the wider context in which interaction takes place. The advent of IUIs is accompanied by a shift of our view of computers, which was facilitated by the transition from mechanistic to linguistic interaction methods (Suchman, 1987). Computers are no longer seen merely as tools, but as potential interlocutors or "social actors" (Nijholt, 2004) and humans tend to attribute to them abilities and traits that they do not have, such as intelligence. Users are willing to interact with artificial entities, such as embodied animated agents, in the same way as they do with other humans and the experience of the former is, to them, similar to the latter (Bickmore, 2004).

The aforementioned findings might justify the view that intelligent systems and environments may be considered as suitable peers for communication, since they allow users to communicate with them in ways that are very similar to human-human communication, thus giving an impression of quasi-humanity. Furthermore, it has been suggested that users who are aware of the possibility that their presence in an AmI environment is being monitored tend to behave differently than if it were not. Thus, an intelligent space may have some of the characteristics of public space (Nijholt, 2004).

Plans and Situated Actions

Planning is often considered the driving force behind actions. Cognitive approaches view plans as the

[3] Implying a combination of environmental circumstances, cognition and the system itself

script for an activity, to be decided beforehand and followed. This model has been favoured by Artificial Intelligence research and models human activity in the formulation and execution of plans (Dourish, 2001). Plans are formulated with a goal in mind. The goal is divided into subgoals, which are in turn divided into lesser goals and so on until a level is reached in which no further division is possible.

The planning model, however, has been challenged by Situated Action theory (Suchman, 1987; Mantovani, 1996): plans do not determine the outcome of an activity; rather, they are fully formed after the activity has taken place in an attempt to describe it or explain the reasons that led to its execution. In short, the structuring of an activity may occur only as a result of the immediacy of the situation (Suchman, 1987; Nardi, 1996a). As Suchman (1987, p.50) puts it, "every course of action depends in essential ways upon its material and social circumstances". Humans often act on impulse and adapt to these circumstances, achieving intelligent action. Consequently, contexts in communication are not preset; rather, they are co-constructed by the participants. Communication is not viewed as the process of information exchange, but as the process of the exchange of meanings and interpretations of the situations the actors are involved in (Riva & Galimberti, 2001).

AmI, by placing emphasis on multimodality, adaptation, and personalisation, is expected to offer computationally viable solutions for the creation of computers that will truly possess the aforementioned qualities, thus overcoming the deficiencies that prevent them from deviating from the planning model and assuming a more human-like behaviour.

Activity theory

According to Activity theorists[4], artefacts such as tools and sign systems assume a crucial role in the shaping of all human experience. As such, activity cannot be understood without an understanding of the role of artefacts in everyday social practice (Nardi, 1996b).

The interplay between action and goals is highlighted by the hierarchical structure of activity (Kaptelinin, 1996). This structure consists of three distinct but interconnected levels. An activity sits at the top level and acts as the direct answer to a specific objective of a subject. This objective characterizes the activity as a whole and is formed in order to satisfy a need. Activities consist of a sequence of actions, which correspond to goals. The fulfilment of actions and goals brings the subject closer to the accomplishment of the objective5. Actions are undertaken by the completion of operations, which reside at the lowest level of the activity hierarchical structure and refer to actions performed automatically, without requiring conscious effort6. Operations are regulated by conditions, which encompass any characteristics of a given artefact that may influence the outcome of the operation.

Certain actions may, due to habituation, become operations: the subject becomes so acquainted with them that conscious effort is no longer necessary for their successful completion. Operations typically move to the domain of actions if the conditions that regulate them change, as for instance in the case of equipment malfunction. In this case, the subject is forced to expend more cognitive resources in order to complete these actions successfully.

To further illustrate this hierarchy, Riva (2005, p.21) offers the example of a person (subject) who wishes to become a psychologist (objective). This might be accomplished by obtaining a PhD in psychology (activity). For this to occur, certain goals will have to be fulfilled first, such as the identification of key sources and the writing of the thesis. These goals require that certain actions be taken, such as going to the library or preparing an index, etc. Of course, one has to employ some form of locomotion in order to go to the library, and needs to type the text in order to submit a dissertation. Walking and typing

4 most notably Vygotsky and Leont'ev

5 and consequently, the activity closer to resolution.

6 This distinction of actions and operations poses some similarities with the distinction between cognitive conscious and cognitive unconscious.

are operations and the conditions, which influence them, are, among others, the location of the library and the position of the keys. Normally, these operations are performed automatically. If, however, the location of the library changes or the subject is presented with an unusual keyboard layout, conscious effort will be necessary.

Similarly, an AmI system is dependent on the physical environmental setting, elements of which are the intelligent artefacts embedded in real space. As the library location and keyboard layout in the above example showed, the arrangement of real space and the artefacts in it may facilitate or impede this procedure, depending on whether it effectively supports the user's spatial reasoning skills. As Suchman (1987) notes, artefacts that are self-explanatory[7] assist in the resolution of a paradox that accompanies modern technologies, namely the requirement of increasingly complex technology being usable with decreasing amounts of user training. The "intelligent" behaviour of artefacts and their intelligibility, would contribute even more towards successful support of users' tasks in the context of real space.

As the aforementioned hierarchy shows, Activity Theory does endorse a somewhat looser notion of planning, at least to the extent that objectives and goals govern one's behaviour. This notion of planning, however, is not as rigid as the planning model described earlier.

Activity Theory demonstrates a more "human-centred" outlook of the relationship between artefacts and human activity in that it does not attempt to portray humans and artefacts as equally important. Rather, human ability is complemented and/or augmented by these external components. Thus, functional organs are formed (Kaptelinin, 1996) which allow humans to gain the ability to perform a previously impossible function or to perform an already existing function more efficiently. In the context of this view, the computer is a type of functional organ that mediates the user's interaction with the environment. This view has led Kaptelinin (1996, p.111) to suggest that we should consider two interfaces, one between the user and the computer and another one between the computer and the environment. In an AmI system, however, the computer is an essential part of the environment (and vice versa) that should facilitate the decomposition of the user's objectives into actions and operations by presenting the user with appropriate affordances[8]. It should also ensure that the conditions of an operation remain largely unchanged, so that the task may be accomplished with a minimum of cognitive effort. To that end, the system should induce a minimum number of user interruptions (breakdowns) and only for the correction of significant problems or the overall improvement of the flow of an activity (Riva, 2005). Support of the user's modalities, ideally in the user's physical environment, helps preserve intuitiveness, thus facilitating habituation and reducing cognitive load.

[7] their purpose and affordance becomes apparent upon examination

Hypothesising about the spatial experience afforded by AmI environments

Merleau-Ponty (1962, p.252) has suggested that "being is synonymous with being situated" and that "… space is existential; we might just as well have said that existence is spatial" (1962, pp.293-4). Thiel (1961, p.35) defines the spatial experience as "a biological function, necessary for the continual adaptation of any organism to its environment, for the purposes of survival". Accordingly, Norberg-Schulz (1971, p.9) explains the spatial experience of humans: *"Most of man's actions comprise a spatial character, in a sense that objects of orientation are distributed according to spatial relations… Man orients to objects; he adapts physiologically and technologically to physical*

[8] An affordance is "a three-way relationship between the environment, the organism and an activity. It is a property of the environment that affords action to suitably equipped organisms (Dourish, 2001).

things… his cognitive or affective orientation to different objects aims at establishing a dynamic equilibrium between him and his environment". These physical objects, which are distributed in space, actually allow for space and the spatial experience as such, by virtue of their formal characteristics and position in space.

Ambient Intelligence (AmI) technology offers the possibility of embedding networked intelligent artefacts in real environments, thus inserting a layer of information with which users interact in the context of an enhanced environmental experience. This experience is no longer only determined by static spatial elements but also consists of intelligent artefacts, accepting as input and displaying multisensory information. Human-computer interaction acquires such characteristics as *disappearance*, *implicitness* and *transparency* and it can be suggested that the interface is spatialised and user interaction now takes place in the context of a real environment. This results in a transformed type of spatial experience, where interaction capability and "intelligence" is bestowed upon physical objects that partly determine spaces as well as on spatial entities themselves.

The interaction between the user and the system takes place within a physical (spatial) and a social context. AmI may influence both of these types of context, which, in turn, may have an effect on the users' behaviour. As such, AmI technologies transform space into a quasi-intelligent entity that keeps track of the users' activities within it via multimodal techniques. As earlier suggested, enhancing the multimodal character of an environment may result in a more complete spatial experience. Also, embedding "intelligence" in real environments, combined with the implicitness of interaction afforded by such systems may result in an environmental experience, resembling the experience of a public space. Indeed, the fact that the user is aware of being monitored makes intelligent spaces similar to public spaces in terms of the way the user's behaviour is influenced. The system's adaptation to the user may facilitate the user's habituation to the system's presence, which may ultimately lead users to consider such spaces as private once again.

The study of theories that describe and explain human communication and behaviour in the context of space is a prerequisite for the formulation of a complete theoretical model which will serve as a tool for the design, implementation, and evaluation of AmI systems. Every theoretical approach has advantages and disadvantages that should be taken into account. The reconciliation of following a plan and improvisation is a complex issue, especially when computers (the archetype of rigidity) are involved. It seems likely, however, that the computer's ability of exhibiting spontaneous behaviour, which will contribute to the support of the user's task, is proportional to its inferential capabilities. These capabilities will not only allow an intelligent system to act on impulse, but also to encourage improvisation on the part of the user by facilitating habituation and helping the user find suitable alternatives when plans go awry.

References

Alcañiz M. and Rey B.: 2005, New Technologies for Ambient Intelligence, in G. Riva, F. Vatalaro, F. Davide and M. Alcañiz (eds), Ambient Intelligence: The Evolution of Technology, Communication and Cognition Towards the Future of Human-Computer Interaction, IOS Press, Amsterdam, pp. 3–15.

Bickmore T.W.: 2004, Unspoken Rules of Spoken Interaction, Communications of the ACM, 47(4), pp. 38-44.

Dourish P.: 2001, Where the Action Is: The Foundations of Embodied Interaction, The MIT Press, Cambridge, MA.

Gibson J. J.: 1986, The Ecological Approach to Visual Perception, Lawrence Erlbaum Associates, London (first published in 1979).

Goleman D.: 2004, Daniel Goleman Omnibus: Emotional Intelligence & Working with Emotional Intelligence, Bloomsbury Publishing Plc, London.

Hall E. T.: 1966, The Hidden Dimension, Anchor Books.

Kuivakari S. and Kangas S.: 2005, Pleasure Platforms and

Sensomotoric Interfaces: Notes from a Preliminary Survey of Adaptive User Interface Design in M. Ylä-Kotola, S. Inkinen and H. Isomäki (eds), The Integrated Media Machine: Aspects of Future Interfaces and Cross-Media Culture, University of Lapland, Rovaniemi, pp. 77–99.

Krueger M.: 1993, Environmental Technology: Making the Real World Virtual, Communications of the ACM, 36(7), pp. 36-37.

Kaptelinin V.: 1996, Activity Theory: Implications for Human-Computer Interaction, in B. Nardi (ed), Context and Consciousness: Activity Theory and Human-Computer Interaction, The MIT Press: Cambridge, MA, pp. 103-116.

Mantovani G.: 1996, Social Context in HCI: A New Framework for Mental Models, Cooperation, and Communication, Cognitive Science, 20(2), pp. 237-269.

Merleau-Ponty M.: 1962, The Phenomenology of Perception, Routledge, London.

Nardi B.: 1996a, Studying Context: A Comparison of Activity Theory, Situated Action Models, and Distributed Cognition, in B. Nardi (ed), Context and Consciousness: Activity Theory and Human-Computer Interaction, The MIT Press: Cambridge, MA, pp. 69-102.

Nardi B.: 1996b, Activity Theory and Human-Computer Interaction, in B. Nardi (ed), Context and Consciousness: Activity Theory and Human-Computer Interaction, The MIT Press: Cambridge, MA, pp. 7-16.

Negroponte N.: 1996, Being Digital, Hodder & Stoughton, London.

Nijholt A.: 2004, Where Computers Disappear, Virtual Humans Appear, Computers and Graphics, 28(4), pp. 467-476.

Norberg-Schulz C.: 1971, Existence, Space and Architecture, Praeger Ltd, New York.

Picard R.W.: 1997, Affective Computing, The MIT Press, Cambridge, MA.

Riva G.: 2005, The psychology of Ambient Intelligence: Activity, Situation and Presence, in G. Riva, F. Vatalaro, F. Davide and M. Alcañiz (eds), Ambient Intelligence: The Evolution of Technology, Communication and Cognition Towards the Future of Human-Computer Interaction, IOS Press, Amsterdam, pp. 17–33.

Riva G. and Galimberti, C.: 2001, Virtual Communication: Social Interaction and Identity in an Electronic Environment, in G. Riva and F. Davide (eds), Communication Through Virtual Technology: Identity, Community and Technology in the Internet Age, IOS Press, Amsterdam, pp. 23-46.

Russell D.M., Streitz, N.A. and Winograd, T.: 2005, Building Disappearing Computers, Communications of the ACM, 48(3), pp. 42–48.

Schmidt A.: 2005, Interactive Context-Aware Systems Interacting with Ambient Intelligence, in G. Riva, F. Vatalaro, F. Davide and M. Alcapiz (eds), Ambient Intelligence: The Evolution of Technology, Communication and Cognition Towards the Future of Human-Computer Interaction, IOS Press, Amsterdam, pp. 159–178.

Streitz N. and Nixon P.: 2005, The Disappearing Computer, Communications of the ACM, 48(3), pp. 33–35.

Suchman L.A.: 1987, Plans and Situated Actions: The Problem of Human Machine Communication, Cambridge University Press, Cambridge.

Thiel P: 1961, A Sequence-Experience Notation, Town Planning Review, 32, April 1961.

Weiser M.: 1991, The Computer for the Twenty-First Century, Scientific American, 265(3), pp. 94-104.