ReCIBE, Año 1 No.1, Noviembre 2012

Collaborative Nonverbal Interaction within Virtual Environments

Adriana Peña Pérez Negrón
Computer Science Department
CUCEI – Universidad de Guadalajara
adriana.ppn@red.cucei.udg.mx

Abstract: Current virtual environments are predominantly visual-spatial, which allows their 'inhabitants' the display, either in a conscious or unconscious way, of nonverbal cues during interaction, such as gaze direction, deictic gestures or location. This interchange of nonverbal messages enriches interaction while supports mutual comprehension, fundamental for collaborative work and therefore particularly important in a multiuser virtual environment, that is, a Collaborative Virtual Environment. Different techniques, the media involvement, and automatic detection related collaborative nonverbal interaction are here discussed.

Keywords: Collaborative Virtual Environment, nonverbal communication, collaborative interaction

Interacción Colaborativa No-Verbal en Entornos Virtuales

Resumen: Los actuales entornos virtuales son predominantemente visual-espaciales, lo que permite a sus 'habitantes' el despliegue, ya sea de manera consciente o inconsciente, de señales no verbales durante la interacción, como son la dirección de la mirada, la gesticulación deíctica o la localización. Este intercambio de mensajes no verbales enriquece la interacción mientras ayuda a una comprensión mutua, fundamental para el trabajo colaborativo y por lo tanto particularmente importante en un entorno virtual multiusuario, esto es, un Entorno Virtual Colaborativo. Diferentes

técnicas, la participación de los medio de comunicación, y la detección automática relacionada con la interacción colaborativa no verbal son aquí comentados.

Palabras clave: Entornos Virtuales Colaborativos, comunicación no verbal, interacción colaborativa.

1. Introduction

During interaction, our nonverbal behavior may comprise most of what we do, including paralanguages cues like loudness, tempo, pitch or intonation of the speech (Patterson, 1983). Moreover, the use of certain objects like the decided outfit, or the physical environment when used to communicate something, without explicitly saying it, has traditionally being considered as nonverbal communication (Knapp & Hall, 2010). Therefore, a simple way to describe nonverbal interaction could be by emphasizing what it is not: the interaction effected by other means than the words signification or meaning.

Nonverbal behavior enriches interaction while supports mutual comprehension fundamental for collaboration (Bolinger, 1985).The functions people give to their nonverbal messages during interaction are to repeat, substitute, complement, accent, regulate or even contradict the spoken message ((Knapp & Hall, 2010), which expose its complexity and large extent. Kujanpää and Manninen (2003) created a satellite type model based in social sciences and communication literature of the different forms of nonverbal communication (NVC) elements, which they claim is an exhaustive set, some of the areas of study included in the model are olfatics, occulesics and chronemics.

However, as extend as it might be in real life, the nonverbal behavior in a virtual environment (VE) is clearly constrained by the media where the senses involved are usually just vision and audition, and eventually some constrained touch feedback through haptic devices. A VE as defined by Schroeder (2011) is:

“a computer-generated display that allows or compels the user (or users) to have a feeling of being present in an environment other than the one that they are actually in and or interact with that environment, - in short, ‘being there’”.

Where the computer- generated display can represent either a real life or an imaginary scenario, and it also can be based on only text, 2 or 3 dimensional graphical representations, however most current VEs are primarily visual experiences. Because of their spatial feature, proper for the display of nonverbal cues, in this paper only 3D representations are discussed.

In order to interact with the virtual world the user requires a graphical representation within it, that is, his/her avatar. Broadly defined, any form of the user representation in the VE can be considered his/her avatar, such as the mouse pointer, although not any representation supports transmitting nonverbal communication cues, those avatars which do better, Salem (1964), categorized and characterized in three groups:

1. Abstract, represented by cartoon or animated characters with limited or predefined actions;
2. Realistic with high level of realism, which imply high cost in technology and hardware resources; and
3. Naturalistic, those with a low-level details approach and that can be characterized as humanoid-like avatars that can display some basic humans’ actions or expressions.

Over time, avatars have become more complex creations with animated movements that aid in the expression of the avatar’s personality and supplement various social interactions (Ahn, Fox, & Bailenson, 2012).

The first 3D animations for humanoids were created only by artistic means, sometimes generating a not complete believable effect of the character’s nonverbal expressions, getting what is known as the “uncanny valley”. This phenomenon is the hypothesis for robots described by Mori (1970) as the relation between human likeness and perceived familiarity, where familiarity increases with human likeness until a point is reached at which subtle

differences in appearance and behavior create an unnerving effect (MacDorman, 2005). Even though, there has been little direct scientific investigation on this effect (MacDorman, 2005), the term has been extended to 3D virtual humanoids. A common practice nowadays, to obtain realistic facial expressions and body language in the animated movies, is the mocap or motion capture, which consists on the transfer of them directly from the actor to the virtual character through different techniques.

Nevertheless, in a computer animated movie or video, the user does not interact with the VE, when interaction is involved in the VE, it is referred as Virtual Reality (VR). The most common classification for VR is related to the users' degree of immersion as desktop-VR, augmented reality and immersive VR:

- In desktop-VR the user can interact with the real and the virtual world at the same time. This technology is considered relatively cheap and therefore easier to spread.
- Augmented reality incorporates computer-generated information into the real world supplementing it with virtual objects that appear to coexist in the same space; and
- In immersive VR the user can interact exclusively with the virtual world, such as with the use of a HMD (head mounted display) like the one shown in Figure 1.

Figure 1. Head mounted display (HMD)

The user interaction with the virtual world is composed by four virtual behavioral primitives (Mine, Brook, & Sequin, 1997):

1. Navigation, the displacement of the user in the virtual space and the "cognitive map" he/she builds of it.
2. Selection, the action of pointing to an object.
3. Manipulation, the modification of the state of an object; and
4. System control, the dialogue between the user and the application, usually going through menus.

Still, in the virtual world the user will be able to interact not only with objects but with virtual inhabitants represented in the world in the same way than the user, by an avatar. And then again, because the focus in this paper is the display of nonverbal interaction in VEs, only humanoid virtual inhabitants will be considered, that by having physical body representation can be very helpful in aiding interaction (Imai et al., 2000)

1.1 Virtual Humans

The research with virtual humans has taken two leading fields as Ahn et al., (2012) pointed out: 1) the use of virtual humans to study social interaction and 2) the use of people to create avatars and agents.

VEs have been used in social science studies because they present a number of advantages such as to allow the researcher to create more realistic experimental situations compared to a lab; also in VEs a lot of the users' movements can be tracked; and the exact same stimulus can be replicated over and over (Blascovich, Loomis, Beall, Swinth, & Hoyt, 2001). As a result, a wide variety of social psychological phenomena have been examined in them, including nonverbal behavior (Ahn et al., 2012).

An example of the study of people nonverbal behavior in VEs is the well-known conducted by (Bailenson, Blascovich, Beall, & Loomis, 2003) where several trials were carried out to understand Proxemics –the study of how man unconsciously structures microspace (Hall, 1968)–, in VEs. On them, the participants clearly treated virtual humans in a similar way to actual humans by keeping their real life proxemic behavior.

For the field of studying people to create virtual humans, a distinction has to be made of humanoid figures from those with autonomy and that can interact with the user, which are considered as intelligent virtual agents (IVAs). IVAs are interactive characters that can communicate with humans or with each other using natural human modalities, therefore its creation involves a number of fields such as sociology, psychology, computer science, artificial intelligence, linguistics and cognitive science.

The other way around the Bailenson et at. (2003) study aforementioned can be (Jan & Traum, 2007), where the authors based on the understanding of Proxemics (Hall, 1968) and how people position themselves in different

situations (Kendon, 1990), formulated a number of algorithms to simulate people movements and position during conversations in agents.

Because of its complexity, facial expressions and conversation face movements represent a great challenge when it comes to implement them in an IVA, even without considering the interaction agent-human. An animation system called RUTH (stands of Rutgers University Talking Head) represents an example of how intricate is to animate nonverbal signals in synchrony with speech and lip movements for agents, this is a freely-available cross-platform developed by Doug DeCarlo and Matthew Stone (DeCarlo, Stone, Revilla, & Venditti, 2004) for this purpose.

2. Nonverbal Collaborative Interaction

Now then, in a VE the user can interact with virtual objects and virtual agents, but in a multiuser VE, that is a Collaborative Virtual Environment (CVE), the user can interact also with other users. Churchill and Snowdown (Churchill & Snowdon, 1998) described CVEs as

"…a terrain or digital landscape that can be 'inhabited' or 'populated' by individuals and data, encouraging a sense of shared space or place. Users, in the form of embodiments or avatars, are free to navigate through the space, encountering each other, artifacts and data objects and are free to communicate with each other using verbal and non-verbal communication through visual and auditory channels".

Here, the user's graphical representation acquires other functionalities becoming the user embodiment in the VE mandatory.

The user's avatar in a VE, as mentioned, is its means for interacting and sensing the various attributes of the virtual world (Guye-Vuillème, Capin, Pandzic, Thalmann, & Thalmann, 1998). But in a collaborative situation it performs other important functions such as perception, localization,

identification and visualization of the focus of attention of the other users (Benford, Greenhalgh, Rodden, & Pycock, 2001; Capin, Pandzic, Thalmann, & Thalmann, 1997).

Gerhard and Moore (1998), who defined the user's avatar as "a proxy for the purposes of simplifying and facilitating the process of human communication" attributed to it five potential properties: identity, presence, subordination, authority, and social facilitation, next described:

1. Identity. Avatars provide to the others in the environment to better understand the concept of an underlie person.
2. Presence. They help establishing a feeling of "being there", a form of self-location
3. Subordination. They imply subordination, that is, they are under the direct control of the user, without significant control over their own actions and internal state.
4. Authority. Avatars act with the authority of the user.
5. Social facilitation. By giving a proxy for human communication and by facilitating interaction.

Related to subordination, the control from the user to his/her avatar, that in turn will affect the avatar's display of nonverbal interaction, can be reached by three different approaches (Capin et al., 1997):

1. Directly controlled, with sensors attached to the user;
2. User-guided, when the user guides the avatar defining tasks and movements, usually through a computer peripheral device such as the mouse; and
3. In a semi-autonomous way, where the avatar has an internal state that depends on its goals and its environment, and this state is modified by the user. For example in videogames, the users' avatar animation displaying joy when the user completes a game goal.

As far as nonverbal features are automatically digitized, directly controlled by the user, they should be more revealing and spontaneous; however, even if nonverbal cues are transmitted to the computer by a simple keyboard or a

mouse, they provide significance to communication and resources to understand collaborative interaction.

The rich of nonverbal interaction in a face-to-face situation is not already available in CVEs; succinct metaphors and words are then the means to substitute it when required; although, the users seem to be able to ignore the absence of many nonverbal cues (Schroeder, 2011). From a very broad point of view people will maintain their nonverbal behavior in VEs as similar as in real life, e.g. the study presented in (Steptoe et al., 2008) where two confederates interviewed a participant in an immersive VE and the participants' gazed at the questioner in 66.7% of cases, a frequency comparable to Argyle's time-range of 70-75% that listeners gaze at speakers during dyadic face-to-face conversations (Argyle & Cook, 1976).

The CVE's characteristics make them better suited for a small group of people (two to five) when a spatial task is involved. The task is likely to be one in which people focus their attention on the space and the objects on it, otherwise these systems would not be used in the first place (Schroeder, 2011). In this type of tasks the other person's avatar body will be used for joint orientation and barely on each other's facial expressions, thus they will not need realistic avatars; it will be sufficient to be able to follow the other's movements and gestures (Steed, Spante, Heldal, Axelsson, & Schroeder, 2003). It has being observed that people treat others' avatars very different when they are socializing that when they are working or doing something together in the VE (Heldal, 2007; Roberts, Heldal, Otto, & Wolff, 2006; Schroeder, 2011), same as in real life.

Now well, some nonverbal behavior varies according to social rules and peoples' nationality (Hall, 1952; Watson & Graves, 1966; Watson, 1970) in such a way that people's background might be part of its analysis. But, even if it is truth that NVC changes from one person to another and from one culture to another, it is also truth that it is functional, which means that different functional uses will lead to specific patterns of nonverbal interchange.

Patterson (1983) proposed what he called “nonverbal involvement behaviors” to operationally define the degree of involvement manifested between individuals; and he classified them within specific functions. These functions are:

- to provide information or to regulate interactions –these two are useful to understand isolated behaviors; and
- to express intimacy, to exercise social control, and to facilitate service or task goals –these last three functions useful to understand behavior over the time.

The first two functions are independent of the last three in such a way that a given behavior can be either informational or regulatory and, at the same time, be part of an overall pattern serving to intimacy, social control, or service-task functions.

In particular, the service-task function identifies the bases for impersonal nonverbal involvement with no reflection of anything about a social relationship between the individuals but only a service or task relationship. The most likely type of nonverbal interaction involvement in a collaborative situation where people take care of a task, which will keep to an acceptable extent cultural and personality influence on nonverbal behaviors, although intimacy and social-control functions will also emerge during a collaborative session.

3. Automatic monitor of the User’s Avatar Nonverbal Interaction

Knapp & Hall (2010), differentiated three primary unites in the study of nonverbal communication:

1. The environmental structure and conditions. This category concerns with those elements that impinge on the human relationship but are not directly part of it. Elements of the environment such as the furniture or lighting conditions; and Proxemics (Hall, 1968).
2. The physical characteristics of the communicators, including his/her artifacts such as clothes, hairstyle or jewelry.

3. The various behaviors manifested by the communicators. The body movements and position also known as Kinesics: gestures, posture, touching behavior, facial expressions, eye behavior and vocal behavior.

These same primary units when transferred to a VE bring up some considerations.

The environmental structure and conditions in a computer display are given by the scenario and the virtual objects around. That is what Hall (1968) differentiated as fixed-features, the space organized by unmoving boundaries such as a room, and semi-fixed features, the arrangement of moveable objects such as a chair. When the communication environment is virtual, the objects there are mainly intentionally located in order to enhance the sense of the place and rarely placed by the user, which will mostly carry out simple scenarios. And probably the most significant difference with a VE compared to a real world environment in this regard, is that typically, only the objects that have a purpose for the task or tasks to be carried out within it can be manipulated, and therefore they must be considered as salient during interaction.

In a computerized environment, the physical characteristics of the interactant will be given by the users' avatar, both appearance and body movements. The range in the appearance of the user representation falls in a wide range; some applications allow their users to create their avatars from scratch, others allow building the avatar from a set of them on which the user can select characteristics such as skin color or clothes, and other applications just give to the user an assigned avatar.

When the environment is for social purposes such as Second Life, the most likely is that the avatar can have a wide range of possibilities for the user to personalize it, and it influences how people treat each other (Schroeder, 2011). In a videogame is more probable to find a set of avatars that will go in concurrence with the game purposes. While for a VE with education or training purposes the avatar will probable have a set appearance with maybe a uniform. Typically, in CVEs for research the users' avatars are naturalistic, which means

as aforementioned, they are humanoid-like that display some basic humans' actions or expressions (Salem & Earle, 2000).

The behaviors of communicators relay on the context that in a CVE will be given by its purpose. For example, in a video game, the users' interaction will be controlled by their intention on getting the goals of the game, while in a social VE the participants interaction will be more likely to be directed to those they feel socially attracted. In Table 1 the primary units of study in NVC are related to the constrained factors in VEs.

Nonverbal interaction influential factors	Constrained in a VE by
Environmental conditions	• the fixed-features of the scenario • the semi-fixed features of the scenario
Physical characteristics	the users' avatars • appearance • body movements
Behaviors of communicators	according to the VE purpose

Table 1. Nonverbal interaction in VEs

The user's avatar body movements and positions, in a VE will probably be adjusted to the software and hardware used to create the virtual environment, and the task at hand. Hitherto, avatars have limited body movements and positions, even when they are tracked directly from the user physical movements, e.g. the most common practice in immersive VEs are the head and one hand movements (Wolff, Roberts, Steed, & Otto, 2005). In, Kujanpää & Manninen (Kujanpää & Manninen, 2003) can be found a considerable set of possible elements an avatar can include for transmitting nonverbal behavior.

As a result, only a limited range of nonverbal interaction can be executed and/or automatically extracted from the VE, and interpreted as part of the collaborative interaction during the session, particularly when there is not vocal content interpretation, this had been discussed somewhere else (Peña, 2011). Based on the criteria of being totally recognizable by a computer system, a list

of nonverbal cues that the users' avatars can display in a VE is presented in Table 2 and next described.

NVC area of study	Cue for automatic monitor in the CVE
Paralinguistics	Amount of talk
	Patterns of talking turns
Objects	When they are part of the collaborative interaction
Proxemics	The location
Kinesics	Deictic gestures
	Gaze direction
	Head movements
	Facial expressions
	Body postures
Appearance	Role

Table 2. Nonverbal cues computer recognizable in CVEs.

Amount of talk and patterns of talking-turns. The paralinguistic's features are harder for computer systems to comprehend than human language. However, the branch that studies, not how people talk, but how much they talk and their patterns of talking-turns has been useful for the study of interaction (e.g. Bales, 1970). Talk-silence patterns, frequency, duration and pacing of speech, have provided means for individual differentiations in social interaction, and in relation to collaborating groups, researchers have found for example, that talkative group members seem to be more task dedicated (Knutson, 1960), and more likely to became leaders (Stein & Heller, 1979). If the channel is written text a posted message can be considered as a talking-turn and in oral communication the microphone can be adjusted to detect the user vocalization.

Artifacts manipulationis an object form of NVC −it can be, for instance, the form that takes an answer to a question (Clark & Brennan, 1991; Martínez, 2003). Therefore, within a CVE, participation can also be related to the manipulation of objects in the shared workspace.

Additionally, according to Jermann (2004), a combination of participation in the shared workspace with amount of talk may be used to establish patterns with

regards to division of labor and the strategies to solve the problems, for example, the alternation in dialogue and implementation might reflect a plan-implement-evaluate approach. In consequence, patterns composed of amount of talk and manipulation in the shared workspace could be useful for the analysis of the collaborative interaction within the VE.

Gazes. Gazes usually have a target, which has to be part of the data collection since this target indicates the user's focus of attention. The gaze is an excellent predictor of conversational attention in multiparty conversations (Argyle & Dean, 1965), and the eye direction is a high indicative of a person's focus of attention (Bailenson et al., 2003). Therefore, via the users' avatar gazes it can be inferred if they are paying attention to the current task and/or to which other participants. Through gazes it is possible to oversee if the group maintains the focus on the task; they can be also helpful to measure the degree of participants' involvement in dialogue and implementation.

Deictic Gestures Gestures have narrative –iconic gesture–, and grounding –deictic gesture– functions (Roth & Lawless, 2002); while it can be difficult to automatically distinguish iconic gestures from the very common meaningless gestures people do when they are speaking, deictic gestures can be easily matched to the mouse pointing.

Deictic terms such as: here, there, or that, are interpreted as a result of the communication context, and when the conversation is focused on objects and their identities, they become crucial to identify the objects quickly and securely (Clark & Brennan, 1991). Consequently, deictic gestures, especially those directed to the shared workspace, will be useful to determine whether the users are talking about a particular object.

Proxemics. When people are standing, they tend to form a circle in which they include or exclude other people from the interaction (Scheflen, 1964). Then, when navigation is part of the CVE, the users' proxemic behavior can be easily

retrieved by the computer system indicating peers' inclusion or exclusion of task activities, the creation of subgroups and division of labor.

Head Movements. Head position can provide a very close approximation to eye direction; head position then could be useful to replace gazes retrieval when it is not possible to follow the exact direction of a person's sight (Parkhurst, Law, & Niebur, 2002), in this case they can be treated like gazes.

On the other hand, there are multitudes of head movements during interaction that have to do with the nature, the purpose and the organization of it (Heylen, 2005). The automatic comprehension of head gestures becomes complex because they carry out different functions and/or meaning that depend on the context in which they are produced. Despite this difficulty, there are some semantic head movements that can be distinguished and can be helpful for collaborative interaction analysis accompanied with other nonverbal behaviors, such as the very common nodding to show agreement or comprehension, or the side to side movement to indicate disagreement or incomprehension. Nods and jerks are typical movements involved to provide feedback.

Body Postures. Body postures are movements that spread throughout the body, visibly affecting all parts and usually involving a weight shift (Bartenieff & Davis, 1972), in contrast to gestures that are movements of only a part of the body. This type of nonverbal cues poses a more complex challenge than head movements because there is not yet a clear association between postures and their interpretation (Mota & Picard, 2003). However, for seated people there seems to be some results like: when people are seated around a table, the degree of orientation between the speaker's torso and the listener can show agreement, liking, and loyalty when aligning with him/her (Mehrabian & Friar, 1969) and, a parallel orientation reveals neutral or passive moods (Richmond, McCroskey, & Payne, 1991).

Facial expressions. Through face, people reflect interpersonal attitudes, provide feedback to others' comments, and it is considered the primary source

of information after speech (Knapp & Hall, 2010). As mentioned, in computer-generated environments one of the main issues has been the creation of realistic-looking facial expressions. Most approaches in CVEs for facial expressions use the widely accepted categorization of Ekman (1978) consistent of six universal basic emotions that can accurately be face expressed in all cultures: surprise, anger, fear, happiness, disgust/contempt and sadness.

The most important feature of facial expressions in a task-oriented collaborative interaction might be to convey understanding feedback to the partners, but it represents a complex challenge to transmit them precisely to the VE. In the context of collaboration, it is worth to mention that eye gaze and facial expression are in many cases critical for interpersonal interaction, but bodily movement and gesture are needed for successful instrumental interaction (Schroeder, 2011). As a result, as Schroeder (2011) pointed out, *"perhaps an avatar face with the possibility to express only certain emotions or only certain acknowledgements of the other person's effort will not only be sufficient in the immersive space but superior –because it will reduce the 'cognitive load' in the task".*

4. Conclusions

Nonverbal cues aid mutual comprehension during collaboration, how people use them or adapt themselves to substitute them in VEs while carrying out a task is still an open issue. In this paper how their display possibilities within the boundaries of a VE were discussed, but only under the assumption of a small group of users doing a spatial task and being represented in the environment by naturalistic avatars. This set shrinks when treated as data for a computer to interpret them. Finally, I have to agree with Knapp & Hall (2010) when they pointed out: *"the nonverbal cues sent in the form of computer-generated visuals will challenge the study of nonverbal communication in ways never envisioned".*

References

Ahn, S. J., Fox, J., & Bailenson, J. N. (2012). Avatars. In W. S. Bainbridge (Ed.), Leadership in science and technology: A reference handbook () SAGE Publications.

Argyle, M., & Cook, M. (1976). Gaze and mutual gaze. Cambridge: Cambridge University Press.

Argyle, M., & Dean, J. (1965). Eye contact, distance, and affiliation. Sociometry, 28(1), 289–304.

Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. (2003). Interpersonal distance in immersive virtual environments. Personality and Social Psychology, 29, 819-833.

Bales, R. F. (1970). Personality and interpersonal behavior. New York: Holt.

Bartenieff, I., & Davis, M. (1972). Effort-shape analysis of movement. the unity of expression and function. Body movement: Perspectives in research (). New York: Arno Press.

Blascovich, J., Loomis, J. M., Beall, A. C., Swinth, K., & Hoyt, C. (2001). Immersive virtual environment technology as a methodological tool for social psychology. Psychological Inquiry, 13, 103–124.

Bolinger, D. (1985). Intonation and its parts: Melody in spoken english. London: Edward Arnold.

Brdiczka, O., Maisonnasse, J., & Reignier, P. (2005). Automatic detection of interaction groups. ICMI, Trento, Italy.

Capin, T. K., Pandzic, I. S., Thalmann, N. M., & Thalmann, D. (1997). Realistic avatars and autonomous virtual humans in VLNET networked virtual environments. From Desktop to Webtop: Virtual Environments on the Internet, WWW and Networks, International Conference, Bradford, UK.

Churchill, E. F., & Snowdon, D. (1998). Collaborative virtual environments: An introductory review of issues and systems.Virtual Reality: Research, Development and Applications, 3, 3-15.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine & S. D. Teasley (Eds.), Perspectives on socially shared cognition (pp. 127–149). Hyattsville, MD: American Psychological Association.

DeCarlo, D., Stone, M., Revilla, C., & Venditti, J. J. (2004). Specifying and animating facial signals for discourse in embodied conversational agents. Computer Animation and Virtual Worlds, 15(1), 27–38.

Ekman, P. (1978). Facial expression. Hillsdale, NJ: Erlbaum.

Ford, C. E. (1999). Collaborative construction of task activity: Coordinating multiple resources in a high school physics lab. Research on Language and Social Interaction, 32, 369-408.

Gergle, D., Kraut, R. E., & Fussell, S. R. (2004). Language efficiency and visual technology: Minimizing collaborative effort with visual information. Journal of Language and Social Psychology, (23), 491-517.

Gerhard, M., & Moore, D. (1998). User embodiment in educational CVEs: Towards continuous presence. 4th EATA International Conference on Networking Entities, NETIES98: Networking for the Millennium, West Yorkshire, United Kingdom.

Guye-Vuillème, A., Capin, T. K., Pandzic, I. S., Thalmann, N. M., & Thalmann, D. (1998). Nonverbal communication interface for collaborative virtual environments. Collaborative Virtual Environments, University of Manchester. pp. 105-112.

Hall, E. (1952). The silent language. Garden City, NY: Doubleday.

Hall, E. (1968). Proxemics. Current Anthropology, 9, 83-108.

Heldal, I. (2007). The impact of social interaction on usability for distributed virtual environments. International Journal of Virtual Reality, 6(3), 45-54.

Heylen, D. (2005). Challenges ahead: Head movements and other social acts in conversations. AISB - Social Presence Cues Symposium,

Imai, T., Qui, Z., Behara, S., Tachi, S., Aoyama, T., Johnson, A., et al. (2000). Overcoming timezone differences and time management problem with tele-immersion. 10th Annual Internet Society Conference (INET), Yokohama, Japan.

Jan, D., & Traum, D. (2007). Dynamic movement and positioning of embodied agents in multiparty conversations. ACL 2007 Workshop on Embodied Language Processing, Prague, Czech Republic.

Jermann, P. (2004). Computer support for interaction regulation in collaborative problem-solving. Unpublished University of Genéva, Genéva, Switzerland.

Kendon, A. (1990). Spatial organization in social encounters: The F-formation system., 209–237.

Knapp, M., & Hall, J. (2010). Nonverbal communication in human interaction (7 th ed.). Belmont, CA: Thomson Wadsworth.

Knutson, A. L. (1960). Quiet and vocal groups. Sociometry, 23, 36-49.

Kujanpää, T., & Manninen, T. (2003). Supporting visual elements of non-verbal communication in computer game avatars. Level Up - Digital Games Research Conference, Universiteit Utrecht. pp. 220-233.

MacDorman, K. F. (2005). Androids as an experimental apparatus: Why is there an uncanny valley and can we exploit it? Computers in Human Behavior, 25, 695-710.

Martínez, A. (2003). A model and a method for computational support for CSCL evaluation (in spanish). Unpublished University of Valladolid,

McCowan, I., Gatica-Perez, D., Bengio, S., Lathoud, G., Barnard, M., & Zhang, D. (2005). Automatic analysis of multimodal group actions in meetings. IEEE Trans. on Pattern Analysis and Machine Intelligence, 27, 305–317.

Mehrabian, A., & Friar, J. T. (1969). Encoding of attitude by a seated communicator via posture and position cues via posture and position cues. Journal of Consulting and Clinical Psychology, (33), 330-336.

Mine, M., Brook, F. P., & Sequin, C. H. (1997). Moving objects in space: Exploiting proprioception in virtual-environment interaction. Computer Graphics Proceedings, ACM SIGGRAPH 1997, pp. 19-26.

Mori, M. (1970). Bukimi no tani (the uncanny valley). Energy, 7(4), 33-35.

Mota, S., & Picard, R. W. (2003). Automated posture analysis for detecting learner's interest level. Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction, CVPR, HCI, Madison, WI.

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. Vision Research, 42, 107–123.

Patterson, M. L. (1983). Nonverbal behavior. A functional perspective. New York: Springer-Verlang.

Peña, A. (2011). Monitoring collaboration in virtual environments for learning: A nonverbal communication approach. U.S.A.: Lambert Academic Publishing.

Richmond, V. P., McCroskey, J. C., & Payne, S. K. (1991). Nonverbal behavior in interpersonal relations (2nd Ed. ed.). Englewood Cliffs, New Jersey: Prentice Hall.

Roberts, D., Heldal, I., Otto, O., & Wolff, R. (2006). Factors influencing flow of object focussed collaboration in collaborative virtual environments. Virtual Reality, 10(2), 119-133.

Roth, W. M., & Lawless, D. (2002). When up is down and down is up: Body orientation, proximity and gestures as resources for listeners. Language in Society, 31, 1-28.

Salem, B., & Earle, N. (2000). Designing a non-verbal language for expressive avatars. Collaborative Virtual Environments, San Francisco, CA, USA. pp. 93-101.

Scheflen, A. E. (1964). The significance of posture in communication systems. Psychiatry, 27, 316–331.

Schroeder, R. (2011). In Kirlik A. (Ed.), Being there together: Social interaction in shared virtual environments. New York: Oxford University Press.

Steed, A., Spante, M., Heldal, I., Axelsson, A., & Schroeder, R. (2003). Strangers and friends in caves: An exploratory study of collaboration in networked IPT systems. Proceedings on Interactive 3D Graphics, pp. 51-54. New York: ACM Press

Stein, R. T., & Heller, T. (1979). An empirical analysis of the correlations between leadership status and participation rates reported in the literature. Journal of Personality and Social Psychology, 37, 1993-2002.

Steptoe, W., Wolff, R., Murgia, A., Guimaraes, E., Rae, J., Sharkey, P., et al. (2008). Eyetracking for avatar eye-gaze and interactional analysis in immersive collaborative virtual environments. CSCW 2008 (ACM 2008 Conference on Computer Supported Cooperative Work),CSCW 2008 (ACM 2008 Conference on Computer Supported Cooperative Work), San Diego, CA, USA.

Watson, O. M. (1970). Proxemic behavior: A cross-cultural study. The Hague, Nederlands: Mouton.

Watson, O. M., & Graves, T. D. (1966). Quantitative research in proxemic behavior. American Anthropologist,68, 971-985.

Wolff, R., Roberts, D., Steed, A., & Otto, O. (2005). A review of tele-collaboration technologies with respect to to closely couplemd collaboration. International Journal of Computer Applications in Technology,

Biographical notes:

Dra. Peña Pérez Negrón Rreceived her Ph.D. in Computer Science in 2009 from the Universidad Politécnica de Madrid, Spain. Her main research interest is on the user's avatar display of nonverbal communication in Collaborative Virtual Environments. She is a research professor at the Computer Science Department at the CUCEI of the Universidad de Guadalajara, Mexico.