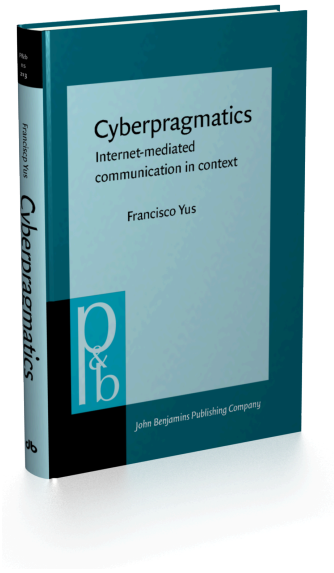
Pragmatics & Beyond New Series
213

5. The virtual conversation

https://doi.org/10.1075/pbns.213.c5



Pages 151–218 of
Cyberpragmatics: Internet-mediated communication in context
Francisco Yus
[Pragmatics & Beyond New Series, 213] 2011. xiv, 353 pp.

CHAPTER 5

The virtual conversation

1. Introduction

On the Internet there are multiple options for engaging in synchronous conversations: chat rooms, instant messaging, Internet-mediated phone calls, videoconferencing, 3D virtual worlds, etc. Some of them have been incorporated into other more typically asynchronous forms of Internet-mediated communication, as happens with the instant messaging application on social networking sites.

In this chapter I will review most of these options for virtual conversations. One of the central issues in this chapter will be to analyse how users compensate for the lack of oral features that their typed texts exhibit, compared to the contextual richness of face-to-face interactions.

2. Chat rooms

The chat room is one of the most popular forms of virtual conversation. Although it has evolved into an enhanced medium with the incorporation of web cams and sound, many users still prefer the traditional text-based utterances sent to a chat portal on the Net. These portals for synchronous conversations contain a number of design features that influence the quality of interpretations and eventual relevance, as will be analysed below.[1] This section deals with the pragmatic implications of chat rooms and of the synchronous oralization of the text typed by users.

Chat rooms are not only interfaces for virtual conversations, but also encourage community bonding. Many of the attributes that were commented upon in Chapter 2 as indicative of a feeling of community among users (Yus 2007b, Baym 2010: 71–98) are also reproduced in these synchronous interactions. Paolillo

---

1. As Baron (2003a) comments, chat rooms, as they are used nowadays, were not created until 1988, when Jarkko Oikarinen, a student at the University of Oulu (Finland), created software that was later known as *Internet Relay Chat* (IRC). At the beginning of the 90s, it turned into an open-access program and started being offered by Internet providers such as *America Online*. For a short history of chat rooms see Mariottini (2004: 29–30).

(2001: 185) asserts that chat rooms can foster the formation and consolidation of social networks that are similar to the ones found in physical settings. Their users spend much time chatting there, which can be compared to "spending some time together" in physical scenarios and these casual conversations foster the formation of networks in urban spaces. Besides, chat room users exhibit linguistic strategies that indicate community membership. This is why non-members are very easy to spot. These "outsiders" will have to "train themselves" in the use of the linguistic strategies that abound in this type of virtual conversation and mark boundaries for group membership. Regular interactions in chat rooms will turn these neophyte users into experts who discursively blend into the community (the *familiarity principle* that Peris et al. 2002: 44 suggest).

2.1 Utterance, propositional attitude and audio-visual context

The steps of interpretation, according to relevance theory (Sperber & Wilson 1986, 1995) start with the identification of the logical form of the utterance, which is enriched inferentially to yield the proposition expressed by the utterance. This proposition can be selected as the explicit interpretation of the utterance (*explicature*) or it can be used as one of the premises that, together with contextual information, enable the derivation of implicated conclusions (*implicatures*). The inference of explicit and implicated interpretations is guided by an inherently human search for relevance (see Wilson 1999, 2000). And the speaker has to predict that the addressee will be able to access the necessary contextual information that allows for the inference of these interpretations. As will be analysed below, in the case of Internet-mediated communication one of these predictions includes the addressee user's command of a number of typical discursive techniques that are inherent in text-based interactions (abbreviations, oralization of text, and emoticons, among others). As Fuller (1994) asserts, if the models that users construct differ greatly, communication may break down: "the models of other people's expectations and prior knowledge that people bring into communication can influence not only the tone of the discussion, but also the expectations of one person regarding someone else's personality." In other words, users need what Simpson (2005) considers as electronic communicative competence for managing and agreeing on conversational strategies and discursive rules that differ so much from oral interactions, and which I label *cyberliteracy*. This explains why chat room users tend to agree on conventions "on the fly" that guarantee mutuality in how certain communicative strategies have to be expressed and interpreted, as in the following example quoted in Campbell & Wickman (2000):

(1)

<Wickmansa>	I was thinking of a coded way to quickly indicate "busy"
<Wickmansa>	that the other person wouldn't respond to
<DRCSC>	ok, what?
<Wickmansa>	but would know to stop sending messages
<Wickmansa>	it could be anything we decide that is one keystroke
<Wickmansa>	like *
<Wickmansa>	or /
<DRCSC>	ok
<Wickmansa>	you pick one
<DRCSC>	doesn't matter
<Wickmansa>	whichever is hit first I guess
	[…]
<DRCSC>	/ may be the easiest
<Wickmansa>	Yes
<DRCSC>	/
<DRCSC>	back
<DRCSC>	that worked

Besides, in the interpretation of the intended interpretation it is important to identify the speaker's propositional attitude upon coding the utterance (or in a more general sense, the relationship between the speaker and the thought expressed by the utterance), because the eventual interpretation will be different depending on whether the speaker is regretting, ordering, asking, advising, etc. with the utterance. This attitude can be communicated in different ways: (a) syntactically (with verbal mood, for example); (b) lexically (using assumption schemas – as in speech acts – that include attitudinal markers such as "I regret that…," "I suppose that…" or "I wish that…," and also with adverbs such as "unfortunately," "probably," etc.); and (c) nonverbally (a smile can reveal an underlying ironic intention). In ordinary interactions, interlocutors devote much effort to identifying the speakers' attitude towards what they are saying and metarepresent their intentions. Some of this cognitive reasoning may be really complex, even though the hearers are normally unaware that they are performing it.

In (2) there is an example of three levels of metarepresentation (summarized in (3)), generated in the hearer's search for the speaker's underlying intention (Sperber 2000):

(2) Mary is picking berries. Peter happens to be watching Mary. Mary intends that Peter should be aware of her intention to inform him that the berries are edible.

(3) Mary intends… that Peter should believe… that Mary intends… that he should believe… that these berries are edible.

These metarepresentational inferences are universal and, like the human search for relevance, they are biologically rooted in human psychology.[2] It is not surprising, then, that chat rooms should reproduce the same inferences of attitudinal attribution, as in this conversation quoted in December (1993):

(4) [*<wabbit> is surprised at a message by <KMOORE>*].

<KMOORE>	wabbit well i thought that you thought that i meant something else!!!
<KMOORE>	wabbit that was a confusing line i just wrote.
<wabbit>	i think what i meant and what you meant didn't mean the same thing and we're all confused now.
<KMOORE>	wabbit yes.. exactly. and now i know what u meant and you know what I meant!
<wabbit>	kmoore: that's what i meant!

As I have already suggested, a textual way of communicating attitudes is to use speech-act expressions such as "I'd like to ask you if…" or "I recommend that…." Curiously, the commands that are used in most chat rooms include keystrokes that automatically turn a sequence of characters into a complete speech act (see Hassell 1998, Cherny 1995b, and Goutsos 2005).

2.2 "What is important is to be able to talk"

2.2.1 *Introduction*

The title of this heading corresponds to a famous slogan of a Spanish phone company in 2001 ("lo importante es poder hablar"), showing the main reason why users spend hours engaged in virtual conversations, namely, to be able to chat with people regardless of their physical location. To achieve this goal, users sometimes have to make the most of typed texts with similar goals to the ones that underlie conversations in physical environments: to attract the interlocutor's attention, direct this attention towards the user's intention and, finally, reveal this intention

---

2. Metarepresentations are essential cognitive operations of humans, according to which, when a person is faced with a mental representation, this person is capable of making a representation of this representation. There are several possible metarepresentations: A thought about another thought, as in (a); an utterance about a thought, as in (b); a thought about an utterance, as in (c); and an utterance about another utterance, as in (d) (see Wilson 1999, 2000):

a. John thinks: Tom *wants me to leave*.
b. Mary says: Tom *thinks that he is intelligent*.
c. John thinks: Mary says that *she ate all the chocolates*.
d. Mary says: John says *that it rains a lot in England*.

(S&W 1986: 153–154). For this triple task, users resort to a number of conversational strategies and alter their texts as much as necessary to convey their intentions, but if these are not properly understood, there may be a significant increase in mental effort or even room for a wrong interpretation.

Therefore, users of chat rooms are expected to master several norms and conventions about how to interact in these environments (Araujo & Melo 2003, Savas 2011: 308). Sometimes users also *exude* information about their command beyond their intention to communicate it explicitly. For example, a user who is sent a message that is full of abbreviations, acronyms, emoticons, etc. may get the impression that the sender is an expert in chat room communication, although the sender may have no intention to communicate that information ostensively (S&W 1986: 58). In fact, there is usually a parallelism between how short and altered the message is and the impression of command that the user exudes.[3] For example, in Chapter 1 the lack of familiarity of <mariabisb> in (5), correlated to a lot of typed text, was compared to the expert user <Bisbaal> in (6a), whose short message communicates much information with just a few keystrokes, as can be seen in (6b), the type of information that the (familiarized) users of this chat room, *Operación Triunfo* (the Spanish equivalent of the British *Fame Academy* programme), can recover without much processing effort:

(5) <mariabisb> rosa tiene una voz bonita pero le falta mucha autoridad en el escenario en eso le dan 100 vueltas chenoa y bisbal, y manu. [*Rosa has a nice voice but she lacks authority on the stage and in that chenoa and bisbal and manu beat her hands down*].

(6) a. <Bisbaal> y creo q n tienen dntro d la academia.

b. <Bisbaal> y[o] creo q[ue] [los concursantes de Operación Triunfo] n[o] tienen [un ordenador conectado a Internet] d[e]ntro d[e] la academia [de O.T.].
[*I think that the contestants at Operación Triunfo don't have a computer logged onto the Net inside the Operación Triunfo Academy*].

Initially, one might think that users interact in text-based chat rooms *despite* the limitations of text-based communication, that is, these users type what they would like to be saying and they read what they would like to be listening to. However, this assertion has to be qualified. In fact, many users interact in these

3. There is a parallel relationship between the level of coherence between messages sent to the chat room and the users' command of the technology to sustain conversations there (see Cornelius & Boos 2003).

environments not "despite" but *precisely because* chat rooms possess this textual quality (Yus 2001b). For example, a user in Savas (2011: 308) explains why he prefers chat rooms: "I am more comfortable with chat because I can overcome my pronunciation problems. When I speak I have to be careful with my pronunciation. In chat, I can express myself better."

2.2.2 *Limitation or advantage*

Anonymity and textual quality are, on first sight, limitations of chat rooms if compared to the richness of face-to-face interactions. Undoubtedly, the vocal and visual information that accompanies speech in situations of physical co-presence are essential for a correct interpretation. "Proximity allows participants to gauge whether or not they are being understood and take appropriate action if they are not" (Freiermuth 2011: 129). Therefore, users are likely to be dissatisfied with some reduction in the possibilities for expressing their thoughts properly on the keyboard and also with the problems involved in checking the extent of each other's cognitive environments. But in reality, many chat room users are satisfied; they hide behind the security of the *nick*, and free themselves from the pressure that being face-to-face with another person exerts. In this environment they express themselves more accurately, freely, spontaneously, and even play with the multiplicity of identities that the system fosters.

According to Caldwell & Taha (1993), it may be concluded that many chat room users shy away from direct face-to-face contact because of the challenge of controlling the interrelation of verbal and nonverbal information. The control that users have over other users' impressions and interpretations may lead to a preference for virtual conversations. But the new developments in this type of interactions on the Net, for example the introduction of web cams and the microphone, add a new dimension because users have to assess to what extent they are willing to let other users perceive their vocal and visual nonverbal behaviour (intentional or exuded) and which impressions they want to convey (see Becker & Stamp 2005, Peter et al. 2007). Nowadays, despite these developments in chat room interfaces, many users still resort to (more secure) plain-text-based communication, except when interacting with close friends and relatives. In that case they are not so reluctant to reveal information.

Altogether, text-based communication is an interesting feature of chat rooms that deserves pragmatic analysis, especially in its oralized quality and in the way it affects processing. Suler (1997a) points out that users find it attractive to see how others express themselves through text in spite of its limitations: "they love to immerse themselves in the quiet flow of words that feels like a more direct, intimate connection between one's mind and the minds of others […] without the distracting sights and sounds of the ftf world" (see also Belson 1994).

2.2.3 *Conversational interaction*

Both in physical and virtual conversations there is a similar goal: to engage in interactions and make information mutually manifest, information that alters each interlocutor's cognitive environment.[4] Chat rooms are "transit places" in which users exhibit their predisposition to interact with one another in a predominantly casual way.[5] Thanks to this interactive environment, users can feel the presence of others and mind-read (i.e. metarepresent) their thoughts, select intended interpretations and agree on the direction that conversations are going to take (Bellamy & Hanewitz 1999). Nevertheless, the differences between face-to-face and chat room interactions are also notorious, as will be commented upon in the next few paragraphs (see Suler 2000).

1. *Synchronous vs. asynchronous*. Face-to-face conversations are centred upon the co-presence of interlocutors and there is a possibility of overlappings and interruptions. Chat rooms are synchronous, but they depend on a rigid succession of messages as they arrive at the computer system (Giese 1998). Although both users who are having a chat-room conversation are online simultaneously, chat rooms suffer from succession in time and space. This fact reduces communicative options and naturalness. And the scenario gets even more complicated if more than one user is involved in a chat room conversation. As a user in Savas (2011:307) explains, "flow and continuity is difficult to establish with more than two people at a time. Continuity can be established with two people if there is substantial "waiting time" while the other person waits for the response from the other. But still, disjointed discussions are common in chat and poor typists suffer." Nevertheless, the evolution in virtual environments for text-based interactions permits nowadays a true feeling of synchronicity. For example, some systems allow users to read, word by word, what the other user is typing on his/her screen, drawing understanding closer to the inferences in face-to-face interactions. In the latter, the processing of previous stretches of the utterance generate anticipatory inferential hypotheses that are (dis)confirmed with the processing of subsequent stretches of discourse. Frequently, listeners have to backtrack and re-interpret previous discourse in the light of new interpretive evidence just processed. This is now possible with modern systems for synchronous text-based interactions (apart from the obvious capacity of voice-enabled chat rooms). As a consequence, classifications of synchronous virtual interactions such as Anderson et al.'s (2010) are welcome:

---

4. Especially now, when both types of interactions tend to be amalgamated in personal networks of a hybrid nature. See Suler (1997c) for a terminological proposal of in-person relationships and cyberspace relationships for physical and virtual interactions, respectively.

5. See, among others, Mayans (2000a, 2000b, 2002a: Chapter 2), Baldwin (1996: Chapter 3), Rafaeli & Sudweeks (1997) and Nilsen (1999).

> Systems can be said to predispose communication to *non-simultaneity* (i.e. asynchronous, one-way systems, such as email, discussion boards, blogs), in which conversation is isolated spatially, chronologically, and contextually; *near simultaneity* (i.e. synchronous, one-way systems, as in instant messaging, multiparticipant chat rooms, and text chat in multiplayer games), in which users typically respond to others' comments as soon as they are received; and *high simultaneity* (i.e. synchronous, two-way systems, such as the VAX "phone" and the contemporaneous UNIX "talk" programs), in which not only immediate responses are possible, but also communications overlap.

2. *Contextual cues in the unfolding of conversations*. The lack of vocal (see 2.3 below) and visual (see 2.4 below) nonverbal information (as part of accessible contextual information) are two challenges for users of text-based chat rooms (Ten Have 2000). To these challenges we can add more technological issues (the *scroll factor*, for instance, the speed at which threads of messages disappear from the top of the screen) and physical ones (there is no shared space, apart from the screen, that anchorages the identification of indexicals), which posit a burden for chat room effectiveness.

3. *Turn taking*. Terms such as move, act, sequence, and opening are typical of conversation analysis. But the most famous term is the speaker's "turn." In chat rooms these turns are subject to the sequencing imposed by the software that manages interactions.[6] Until more "natural" chat room software becomes popular, for instance software (already available) that conveys the user's voice, as in *Second Life* (see 4 below) or one's visual image, the sequencing of utterances will be the norm in chat room interactions, unlike face-to-face interactions. Besides, in chat rooms all utterances are sequenced, even if simultaneously produced. As Nilsen & Mäkitalo (2010: 92) comment, "chat systems are designed so that several persons can post messages simultaneously, which means that there is no competition for the floor since all messages sent off will be posted."

4. *Multiple interactions*. One of the most typical attributes of chat rooms is the juxtaposition of conversations in the same (main) area of the screen, also typical of instant messaging (see 3 below) and *Twitter*. Indeed, unless the user is engaged in a private conversation with another user (in a different window), the norm is that all the messages arrive at the "central area" of the chat room, together with messages that the system creates automatically. This quality might produce increased effort when following threads of conversation that are mixed up without a clear arrangement, and hence affect users' eventual estimations of relevance (Werry 1996: 51).

---

6. See Zitzen & Stein (2004: 991–993). The convergence of interactive aspects of humans and attributes of computer software for Internet communication is usually called *groupware* (Feenberg 1989: 28). The term is also used to describe the creation, among multiple users, of a single text (Greller & Barnes 1993).

Furthermore, while an "addresser user" is waiting for another user's reply, he/she may initiate other conversations in the central area of the chat room or privately with other users in other rooms of the site, or engage in several one-to-one private conversations, and therefore when this user receives the initial addressee user's reply, this might be totally irrelevant for someone who is already carrying out other interactions. This is why Serpentelly (1992) concludes that a "serious" conversation in a chat room is impossible, because we cannot control the multiplicity of simultaneous dialogues that can co-occur in the same virtual space.

A possible strategy of mitigation that reduces increased mental effort when the user tries to follow the conversational threads in the chat room is to type the *nick* of the "addressee user" at the beginning of the message, similar to the convention "@username" in *Twitter*. In this way, it is easier to obtain an adequate level of coherence among the threads. Herring (1999) studied, precisely, the coherence between turns in chat room conversations. Her conclusions were that these conversations suffer from fragmentation, and this provokes a growing lack of interest in users, who have to pay attention to multiple dialogues simultaneously. Herring (ibid.) wonders why chat rooms are so popular if they are so communicatively limited. Two explanations are plausible: (a) the users' ability to adapt themselves to the peculiar idiosyncrasy of chat rooms; and (b) the advantages of losing coherence in exchange for higher interactivity and textual playfulness. In her study, Herring analyses the short conversation in (7), which exhibits an overlapping of conversational threads, as summarized in Table 5.1.[7] There are up to three levels of juxtaposition, and the "jumps" among threads are constant:

(7)

1.	<ashna>	hi jatt
2.	<Dave-G>	kally i was only joking around
3.	<Jatt>	ashna: hello?
4.	<kally>	dave-g it was funny
5.	<ashna>	how are u jatt
6.	<LUCKMAN>	ssa all
7.	<Dave-G>	kally you da woman!
8.	<Jatt>	ashna: do we know eachother?. I'm ok how are you
9.	<kally>	dave-g good stuff:)
10.	<Jatt>	kally: so hows school life, life in geneal, love life, family life?
11.	<ashna>	jatt no we don't know each other, i fine
12.	<Jatt>	ashna: where r ya from?

7. See O'Neill & Martin (2003) and Panyametheekul & Herring (2003) for similar analyses of overlapping threads of conversations in chat rooms.

Table 5.1 Overlapping of conversations in a chat room (adapted from Herring 1999)

User (initials)	**[1] A**	**[2] D**	**[3] J**	**[4] K**	**[5] A**	**[6] L**	**[7] D**	**[8] J**	**[9] K**	**[10] J**	**[11] A**	**[12] J**
Dialogue 1	A		A/B		B/C			C/D			D/E	E
Dialogue 2		A		A/B			B/C		C			
Dialogue 3									A	A		

5. *Ephemeral conversations*. One of the drawbacks of the chat room interface is that the messages keep arriving at the main area and, if there are many users in the same channel and are participating actively, their messages immediately disappear from the top of the screen and it is sometimes impossible to process the information that the users intended to communicate (the so-called *scroll factor*). This affects expectations of relevance and the effort that needs to be devoted to a fast processing of these ephemeral messages, and it even constrains a correct contextualization of these messages.

In general, this fast scroll of messages on the screen has consequences, for example regarding their brevity. It is indeed difficult to find messages of over three lines. Werry (1996: 53) and Rintel & Pittam (1997) suggest another important reason for these short messages: the users are fighting for other users' attention. A potential reader of one's messages may focus his/her attention on another conversation if one spends too long typing the message. In this case, the eventual relevance of one's messages has to be predicted very quickly. In chat rooms, cognitive effects must offset the mental effort required to process the messages that do not stay on the screen for very long. In this sense, Ruane's (1999) words are illustrative:

> It's fast: Try talking to six people at once. It's brief: three or four words per exchange. It takes wit, concentration and nimble fingers. And it requires tremendous linguistic economy. There's neither time nor space for exposition. The solution has been to abbreviate, contract and condense. On a huge scale. Why, for example, consume precious keystrokes telling six friends you have to go smack your little brother when BRB (be right back) will do?

But constant advances in chat room communication and interface allow for an increasingly contextualized and "natural" interaction in this virtual environment.[8] Advances have brought voice and image to the chat room. But if one likes to preserve anonymity, there are also 2D graphic chat rooms such as The Palace (www.thepalace.com), where users can choose a virtual image (which, in Section 4, will be labelled *graphic avatar*) that complements their *nicks*. This is a kind of ritual in which we can wear masks for performing interactive tasks (Goffman 1987).

---

8. See, among others, Suler (1997b, 1999), Smith et al. (2000), and Smith, Cadiz & Burkhalter (2000).

6. *Clipped messages*. Another aspect of chat rooms that may influence message comprehension is that very often a message turns up on the screen divided into chunks that are allocated by the server as different turns and in different positions on the list of messages. For example, in (8), taken from a corpus of chat room conversations (Yus 2003b), the answer to utterance 262 is divided into two messages (268/271) and between them there are messages by other users. The counter-reply (285) can only be found after several messages. Finally the initial user sends another two messages (292/293), this time luckily allocated by the server contiguously:

(8) 262 <fryski> ves soys unos pardillos
[*You see? You are novices*].
268 <zen80> eeeeeeeeeeeeeen fryski no te pases
[*fryski don't push us too far*].
271 <zen80> que de pardillos aki nadie
[*there's no novice here*].
285 <fryski> q pasa zen no te pongas nervioso
[*what's up zen, don't get nervous*].
292 <zen80> no fryski nervioso nooooooooooooo
[*no frysky not a male nervous*].
293 <zen80> nerviosa en todo caso
[*a female nervous, if any*].

7. *Conversations in progress*. One always enters a chat room and finds threads of conversations that are already "in progress" (Mayans 2002a: 34). The user is faced with many messages, some addressed to a general audience and some to specific users, and has to try to make sense of them and decide in which thread he/she is willing to participate (Yus 2009e). In this sense, Miura & Shinohara (2005) propose a model of interpretation of chat room messages that starts, precisely, at the moment when the user accesses the main area of the site. At a first stage, the *information-acquisition phase*, the users face a screen filled with conversations that have already been initiated. They have to structure these messages in coherent threads, understand the context in which they were typed, and devise their own messages. Since messages tend to disappear very quickly from the screen, this task has to be performed very rapidly.[9] Some of these messages and

9. Therefore, although in theory typing on the keyboard gives users time to devise their messages, in reality these are as spontaneous and unplanned as oral utterances in face-to-face conversations. I do not agree with Pano (2008a: 89) when she states that the action of typing characters one by one and sending them to the server to appear on the screen reduces spontaneity and gives users a longer time to think about what they actually mean to communicate, compared to what one really says in a prototypical oral conversation, full of slips.

conversations are held in the user's short-memory storage, including the topics that these conversations are about. This leads to a *situational awareness*. Then the users have to decide in which conversations they are going to participate. At the second stage, called *information output*, the users decide what message they are going to type. To do so, the users devote more processing effort to interpreting some of the messages, and the information they communicate will probably be remembered longer (and stored in the long-term memory). At the same time, this memory is accessed in the phase of composing the messages, together with information on how to use the chat room interface and its typical discursive techniques of oralization.

2.3 Compensating for the loss of the audio channel in chat rooms

2.3.1 *Introduction*

Virtual conversations do not differ so much from oral conversations in terms of what steps interlocutors have to follow in order to reach a relevant interpretation, but they do differ in how these steps are performed. As in any conversation, chat room users choose, among a range of possible options for coding their messages, the one that adequately serves as evidence of their underlying intentions and makes interlocutors grasp the intended interpretation without increased mental effort. Among the predictions, the "addresser users" will expect interlocutors to know about typing conventions and how to oralize text (abbreviations, acronyms, emoticons, among others) and about the use of chat room software (how to send a private message, how to change *nicks*, etc.). This is the type of information that users expect to belong to their mutual cognitive environment. In (9), the author of message 1 presupposes that the readers know the kind of message it is and how to type it, and the same happens in message 2, where it is presupposed that the user knows how to access a channel of the chat room and that the nomenclature "#" is necessary. Finally, in 3–4, we can see an adjacency pair in which a user asks how to type a private message and another user replies:

(9) 1 <ESIGUAL> Porfavor no mas privados
[*please no more private messages*].
2 <ESIGUAL> el chat con naim es en #naim_thomas
[*the chat with naim is in #naim_thomas*].
3 <Naiara> como se hace un privado???
[*how is a private message typed?*].
4 <ESIGUAL> Naiara- /q NICK

Therefore, chat rooms demand some kind of *cyber-literacy* as to how to oralize text and how to use the different commands for engaging in virtual conversations. The former is not only typical of chat rooms, but also of instant messaging and even dialogues in social networking sites. As will be analysed in 2.5, chat room messages are *oralized written texts*, hybrids of typed texts and the users' willingness to communicate their thoughts orally. It is even possible that users "hear" their own voices while they are typing their messages. This "written voice" leads to a textual deformation that aims at transcribing on the screen the message that the speaker would have said orally in a face-to-face conversation. This idea is corroborated in Savas (2011), where a number of informants who exhibited forms of text oralization in their messages were asked if they perceived chat as a written or spoken form of language, and most of them considered their chat room discourse to be "talking." For these informants, "chat was a different way of having a conversation. The only difference between online chat and spoken language was the typed responses in chat. Their choice of vocabulary, grammar, and style during the online discussion reflected how they spoke rather than how they wrote" (ibid.: 309).

This strategy of oralization as a hybrid form of Internet-mediated communication involves a number of techniques that will be studied in 2.5 below. These may strike the reader as unusual or, if they have become conventionalized among the regular chat room users, they might even go unnoticed as a taken-for-granted code. Indeed, textual deformation in chat rooms is continuously contrasted with well-established, normalised forms of written communication against which the users of chat rooms rebel. In Kataoka's (2003: 125) words,

> graphemic features may serve as a means of the writer's affiliation with particular groups, community, contexts, and cultures. Affective signs, exploited by young writers with a certain emotional drive, can index facets of the encoder's self through the ways s/he reveals and responds to affective events. We could take affective signs and punctuation to serve as a means of connecting emotion and youth identities […] Youth identities are closely tied to the community-sanctioned ways of representing emotions that are shared between senders and addressees and appropriate to the epistolary context.

2.3.2 *Typographic innovation*

The keyboard offers users some possibilities for connoting typed text (Reid 1994: 31–32, Ruedenberg et al. 1994, Mariottini 2004). Jaffe et al. (1995) group these under the generic label of *emotexts*, a wide-ranging term that includes intentional variations in orthography (e.g. visual dialect), strategic use of capitalization, lexical substitutions (metalinguistic cues of paraverbal quality, for example

to type "hmmmmm"), grammatical markers (e.g. reiterative use of exclamations), and iconic compositions of characters (emoticons). These possibilities qualify chat rooms as purposely playful.

In chat rooms, a frequent typographic resource for vocal connotation of text (e.g. to give emphasis) is phonematic repetition, as in (10), while other vocal qualities such as "shouting" are communicated with the aid of capital letters, as in (11):

(10) <tardío> tenia gaaaaaaaaaaaaaaaaaaaanas de verte! dooooooooonde estabas?
[*I was looking forward to seeing you! where were you?*].

(11) <tardío> NO LO SE!!!! YA TE LO HE DICHO!!!!! alguna otra pregunta?
[*I don't know! I told you! Any other questions?*].

It should be borne in mind that there are techniques of typographical compensation that are inherent in each language. Since Internet allows for multi-geographic synchronous conversations, a process of mutual awareness of (and eventual agreement on) conventions often has to take place to avoid misunderstandings (Fouser et al. 2000, Launspach 2000). This risk of misunderstanding depends on the existence of mutuality and whether users are strangers or already know one another offline. As Kavanaugh et al. (2005) correctly stress,

> While there is still a need for such verbal cues online in networked geographic communities, the fact that most people already know each other at least as acquaintances, provides a lot of background information about personality and manner from prior face-to-face interactions […] The fact that members of groups interacting online typically already know each other in networked communities mitigates against some of the problems of social presence online.

2.4 Compensating for the loss of the visual channel in chat rooms

2.4.1 *Introduction*

In text-based chat rooms there is no simultaneity between the user's verbal and nonverbal behaviour. The hypothetical nonverbal behaviour that would take place while the user is typing also has to be typed (Reid 1994: 21–32). The user who reads the message has to imagine what nonverbal behaviour accompanies it. Instead of "reading" the nonverbal contextual cues that the speaker *exudes*, the reader has to make do with the *spectrality* (Feenberg 1989: 25) of their mere textual manifestation.

Of course, many authors find a positive side to this lack of nonverbal information, the most obvious being the suppression of stereotypes and social prejudices associated with the user's personal image (Lameiro & Sánchez 1998, Lee 1996). Walker & Bakopoulos (2005) also value positively this "visual silence" that

suppresses the social obsession with visual appearance and liberates us from the limitations that it imposes on us. The reader, then, has to hypothesize this absent nonverbal information (Suler 1997c), which often does not match the interpretation that it would have produced in a face-to-face scenario. A quote by a user (in Baldwin 1996: Chapter 3) is illustrative about this point:

> I was challenged in the sense that I was having to interact with people who had no faces, including no smiles or frowns, people without hands and bodies so that I might read their body language […] I did not know how they looked, how they dressed, nor was I able to hear their distinct southern drawls or haughty yankee accents; I was forced in this manner to be totally objective. They were merely words on the screen.

2.4.2 *Nonverbal communication and simultaneity with verbal language*

An important aspect of text-based chat room language is that verbal and nonverbal communication do not overlap or occur simultaneously, but can only be typed in succession. In face-to-face interactions, the intersection of (non)-intentional nonverbal behaviour and verbal utterances can generate four main prototypical cases (Yus 1997b: 334–335):

1. A non-intentional nonverbal behaviour strengthens the information provided by the verbal utterance.
   [*a person shivers and his hands sweat while saying "I'm nervous"*].
2. A non-intentional nonverbal behaviour contradicts the information provided by the verbal utterance.
   [*a person shivers and his hands sweat while saying "I'm not nervous"*].
3. An intentional nonverbal behaviour strengthens the information provided by the verbal utterance.
   [*a person puts his index finger on his mouth while saying "Be quiet!"*].
4. An intentional nonverbal behaviour contradicts the information provided by the verbal utterance.
   [*a person winks ostensively while saying an ironic remark*].

These possibilities of overlapping are absent in text-based chat rooms because of the succession of typed characters on the screen, a problem shared by all written texts. Hence, it is not correct to claim, as Danet et al. (1998) or Ruedenberg et al. (1994) do, that the interpretation of both sources of information is simultaneous (a kind of *gestalt* following the terminology of Danet et al. ibid.).

2.4.3 *Intentionality in verbal-nonverbal information*

In text-based chat rooms there is no unintentional nonverbal behaviour, that is, information conveyed nonverbally that the users *exude* without consciously intending to communicate it, because all written texts involve a willingness to

produce them (Masterson 1996, Lang 1995). By contrast, it is possible to misunderstand the intentional nonverbal behaviour typed by the user. The two main strategies of intentional transference of nonverbal information are the *emoticon* (2.4.4 below) and the *stage direction* (2.4.5 below).

2.4.4 *The emoticon*

This is one of the most famous ways of connoting typed text with visual nonverbal information (also called *smiley*) (see Quinion 1996, Witmer & Katzman 1997). The most typical ones are "happiness" [:-)], "sadness" [:-( ] and "wink" [;-)]. They were initially inherent in chat room communication but have been exported to instant messaging, mobile phone *texting*, *Twitter* and social networking sites, among others. The source of the term, a combination of the words *emotion* and *icon*, is misleading. As Dresner & Herring (2010: 252) correctly clarify, the purpose of emoticons is not only to express the users' emotions, since

> many facial emoticons do not seem to express a single emotion, or indeed any emotion at all. Is a face with the tongue sticking out – for example ;-p – a sign of a specific emotion? Various sources attribute to it the meanings of teasing, flirting, and sarcasm, all of which may be associated with emotional states, but are not emotions per se. Or consider the familiar winking face ;-): Conventionally, it indicates that the writer is joking, but surely jokes are not associated with a single emotive state. People may joke when they are happy or sad.

In some contexts (institutional ones, for instance) the use of emoticons is considered inappropriate, but they are common in many *cyber-media* and signal that the user is aware of the typical strategies for oralization.[10] Nowadays, the production of emoticons has been automatized by the software and even 3D and Flash-animated emoticons populate chat room conversations turning them into a kind of *pastiche* of colourful verbal-visual information. Besides, the current software for chat room management offers the option to generate emoticons by typing a sequence of characters. For example, by typing ":," "-" and ")" the system immediately generates a 3D emoticon of smile.[11]

---

10. Emoticons are typically informal, so it is not surprising that Derks et al. (2007) concluded that the use of emoticons is much more frequent in environments of camaraderie, friendship or tight-knit groups than in more neutral scenarios such as the workplace.

11. The emoticon was created in 1982 by Scott Fahlman, from Carnegie Mellon University. At that time, Fahlman wrote: "I propose the following character sequence for joke markers: :-). Read it sideways. Actually, it is probably more economical to mark things that are NOT jokes, given current trends. For this use :-(." (quoted in Baron 2003a).

In this book, the term *emoticon* will refer to textual combinations of characters to create iconic images. However, other authors propose a wider semantic field for this term. It is the case of Metz (1994: 41), who proposes four types of emoticon: (1) verbalizing nonverbal behaviours (I will call this *commented stage direction* in 2.4.5 below); (2) writing nonverbal behaviours between asterisks (I will call this *autonomous stage direction* in 2.4.5 below); (3) marking the text with capitalization; and (4) composing images with punctuation marks (the only type that, in my opinion, should be labelled "emoticon").

Furthermore, emoticons are intentional and do not cover the whole range of nonverbal behaviours that people give off or *exude* without intention. In an ordinary situation such as (12), B's nonverbal behaviour – a yawn – can lead to many interpretive possibilities (13a–f) depending on the (in)adequate identification of the underlying intentionality (or lack of it) and on the correct interpretation of this nonverbal behaviour. But in text-based chat rooms only (13a) and (13c) are possible:

(12) [*During A's visit, the host – B – yawns*].

(13) a. B yawns intentionally so that A infers that B wants him to leave; A interprets the ostensive act correctly and decides to leave.
b. B yawns intentionally so that A infers that B wants him to leave; A interprets the ostensive act as a signal that B is tired, that is, A thinks that there is no underlying intention in the yawning.
c. B yawns intentionally so that A infers that B wants him to leave; A correctly interprets the act as ostensive, but chooses a different interpretation, for example that B is trying to communicate that A is very boring.
d. B yawns because she is tired, without any intention in its production; A interprets correctly both the behaviour as unintentional and its interpretation (tiredness).
e. B yawns because she is tired, without any intention in its production; A interprets correctly the behaviour as unintentional but chooses a different interpretation, for example boredom instead of tiredness.
f. B yawns because she is tired, without any intention in its production; A incorrectly interprets the act as ostensive and connotes it with an intentionality that B does not hold.

The origin of these iconic compositions of characters lies in the limitations of traditional punctuation marks that prevent authors from expressing vocal and visual qualities of conversations (Yus 1998g: 119 ff). On the Internet, the plain ASCII text, developed for the *global village* so that all computers with different operating systems could exchange messages, is also a hindrance to users' expressiveness, so they devised new conventions for connoting texts and bringing them closer to their communicative goals.

However, the analysis of emoticons reveals that they are frequently redundant with respect to the typed text that they precede, follow or are inserted within:

(14) a. I am very happy today :-)
b. Yesterday I had a terrible day :-(

Derks et al. (2008) concluded that using an emoticon produces positive judgements among users because it adds a supplementary emphasis compared to the neutral interpretation that the message would convey without it. For instance, the author of a positive message accompanied by the typical "smile emoticon" obtained an interpretation of more happiness with the message.

Additionally, the use of emoticons is reminiscent of the strategy of *lettering* in comics (Gasca & Gubern 1988, Gubern 1992, Yus 1997b: III.1, 2008f). This is a technique for adding emphasis to text in comics balloons by changing, for example, the size and shape of letters, and thus providing readers with a more accurate picture of what the character is feeling.

Although the main function of emoticons is redundancy, sometimes they can alter the meaning of the message, for example reducing its force or even invalidating its propositional content altogether, and on these occasions they play a similar part to nonverbal behaviour in face-to-face interactions. In these oral exchanges, a wink by the speaker can neutralize the effect of the propositional content of the utterance and turn it into an ironic remark, or connote it with politeness (Wilson 1993, Menges 1996). Similarly, an emoticon can soften the meaning of a message and even make it mean the opposite of its literal meaning, as happens in the "wink" that connotes the message (15) with irony:

(15) Only an idiot like you would have done something like that ;-)

Sarcasm would also be conveyed in a similar fashion. The previous example illustrates how an emoticon can aid interlocutors in finding an interpretation that differs from the content of the message completely, which contrasts with Derks et al.'s (2008) opinion that emoticons do not have the strength to turn around the valence of a verbal message. According to Dresner & Herring (2010: 253), these emoticons "seem to have no self-standing content on their own, but rather contribute to – indeed, provide a vital cue as to how to interpret – the linguistic content of messages. When used this way, emoticons seem to be a part of the text, as much as punctuation marks, which can also signal sarcasm." Besides, these authors propose that emoticons have a role beyond redundancy, namely,

> indications of the illocutionary force of the textual utterances that they accompany. As such, they help convey the speech act performed through the production of the utterance. These uses of emoticons neither contribute to the propositional

> content (the locution) of the language used nor are they just an extralinguistic communication channel indicating emotion. Rather, they help convey an important aspect of the linguistic utterance they are attached to: What the user intends by what he or she types. (ibid.: 255–256)

The problem with a non-redundant use of emoticons is that these combinations of characters may become really complex compositions and the reader may not grasp their meaning correctly. Indeed, even in the use of emoticons there is a continuum from the purely arbitrary to the explicitly iconic based on cultural rules of making sense of visual signs. These rules are shared by the community, that is, mutually manifest to all its members. In the case of the emoticon, it demands *emoticonic literacy* from the users beyond the simplest well-known compositions (Reid 1994: 31–32, Watson 1996). Therefore, it is very likely that the authors of the emoticons in (16) will not obtain the desired effect without the aid of the information of the message that precedes them, which *anchors* their meanings, in Barthes' (1977) sense, reducing the range of possible interpretations that these emoticons can convey:

(16) a. You know I don't agree :-e
b. I've recently been ill with a cold :-'I
c. Did you miss me? >:-> (malicious comment)
d. I send you lots of kisses :-X

More examples are found in the following messages (del-Teso-Craviotto 2008: 260): in (17a), the text that follows the two emoticons *anchors* their meanings as "cup" and "rose." In (17b), the user's *nick* helps us to deduce that the emoticon portrays a cowboy winking and his corresponding hat:

(17) a. TNCharmer: c(_) @}}~~~ coffee and roses for the ladies and hello room
b. BGHEARTEDCOWBOY: single m with pic on profile c);o)

Both in redundant emoticons and in cases where emoticons play a more important role, users seem to infer that the emoticon influences the complete message. According to a study by Provine et al. (2007), users tend to process the message as a whole and *then* add the meaning of the emoticon as additional or complementary information. After a classification of emoticons into (a) those which constitute the only content of the message or *naked emoticons*, (b) emoticons that are placed at the beginning or the end of the message, and (c) emoticons that are inside the message, they concluded that the second type is much more frequent. This corroborates a tendency to use emoticons as qualifiers of the whole message.

A possible solution to this inevitable shift of the emoticon to the more arbitrary side of the aforementioned continuum is to find some form of conventionalization, that is, to get all users to know on which occasions all types of emoticon

can be used and with what meaning, independently of the message that they qualify. In this sense, there have been multiple attempts to establish a code among Internet users, especially from glossaries on websites, but these were doomed to failure. Faced with the lack of semiotic code for the use of emoticons (beyond the universal "smile" or "wink"), users are forced to introduce their most original iconic compositions only through redundancy towards the verbal content of the message, for fear of being misunderstood:

(18) a. I like chocolate a lot, :-P
b. I have recently changed my hairstyle &:-)
c. I went for a drink last night #-)
d. Now I have my hair parted in the middle {:-)[12]

The lack of a shared code for emoticon use is corroborated by the fact that the same emoticon is defined differently even in the dictionaries and glossaries that compile them. In Belda Medina (2000: 573, 2003b) there is a comparison between the definitions suggested by Silverlink (a), Milner & Burrows (b) and A. Fernández (c):[13]

(19) :-P (a) stick out one's tongue; (b) to talk in a joking mood; (c) to howl.
#-) (a) to be dead; (c) to go partying all night.
:-7 (b) to smile in an unnatural way; (c) to smoke a pipe.
:-3 (a) to eat a lemon; (c) smile of a man with a moustache.

Sometimes, the emoticon may even be the protagonist of all the interaction, and the knowledge and transfer of information through these emoticons provide users with a playful atmosphere and awareness of mutual manifestness of emoticon conventions. The following conversation in (20), quoted in Merchant (2001: 301), is an example in which most of the information communicated is emoticon-centred:

---

12. Most emoticons are interpreted by turning one's head 90º to the left. However, there are also horizontal emoticons, especially in Asian cultures. In the corpus of *tweets* analysed in Chapter 4, some of these emoticons can also be found, such as [Ò.Ó] and [ò.ó]. Similarly, Baron (2009) comments that American emoticons are read sideways and emphasize the mouth, whereas the Japanese *kaomoji* are read horizontally and focus on the eyes. For example, the typical emoticon for smile, :-), is the *kaomoji* ^--^. Several studies indicate that cultural differences between Japan and the US are reflected in the ways the Japanese interpret these two forms of nonverbal online expression.

13. (a) Silverlink: "Acronym list (including smilies)," (b) A. Milner & T. Burrows (1997): *Internet*. London: Dorling Kindersley; and (c) A. Fernández Conde-Cuadra: "Expresiones en el chat."

(20)
adz46:	hows you
pintsize:	fine thanx u?
adz46:	great
pintsize:	cool wot u up2?
adz46:	not A LOT
pintsize:	wot av u bin up2?
adz46:	Writeing a Macbeth Essay
pintsize:	o gr8 fun!
adz46:	mmmmmmm
adz46:	:-(
pintsize:	:)
adz46:	:-(
pintsize:	cheer up!
adz46:	:-|
pintsize:	Stop it!

2.4.5 *The stage direction*

With this term, I refer to texts that users type in order to describe nonverbal behaviours. It is taken from theatrical terminology, due to the similarities with stage directions that are typical in plays. I distinguish two varieties of stage direction:

1. *Commented stage direction*. In this variety, the text describes, as a whole sentence, the user's nonverbal behaviour:

(21) Tom is laughing out loud.

This is also labelled *emote* in some studies (e.g. Herring forthcoming). Commented stage directions are used not only as verbalizations of nonverbal behaviour, but also with the aim of creating an atmosphere or contextual environment that colours interactions with a higher feeling of realism. For example, in Gelléri (1998) some of these stage directions can be found:

(22)
a. Deadcow waves to everyone
b. neichy1 waves goodnight to jazzzz
c. Ik4u laughs at dinorex
d. frankay is happy now :)

The third person alleviates the monotony of the first person in chat rooms, turning conversations into a kind of theatrical script. This explains why it is not uncommon to find examples of stage directions that refer to the (unseen) environment (Gelléri ibid.):

(23) a. SteveC turns his central heating up, mmmmmm nice
b. DaProphet clears his throat
c. ^Prophet^ is pissed off coz his computer hanged
d. Merlyn was on the phone

These stage directions may comment on events, places or situations that have little to do with the on-going conversation. Sometimes, though, these do reflect moods and feelings that play a part in the interaction (Gelléri ibid.):

(24) a. Kali yawns
b. Graeling sits in the corner and hopes someone will talk to her

In this sense, Cherny (1995b) proposes five types of commented stage direction, exemplified in (25a–e):

(25) a. lynn waves.
b. lynn nods.
c. Mike pastes Tom's lips…
d. lynn packs for the trip.
e. lynn hated the film.

(25a) is a *conventional action* typical of chat rooms, waving upon entering the room. (25b) is a *back channel*, that is, a discursive strategy to make manifest the user's interest in the current conversation. (25c) is a *meta-discourse comment* of a humorous kind about another user's message. (25d) is a prototypical commented stage direction, specifically a *narration* of an ordinary nonverbal activity transcribed as text. Finally, (25e) is an *exposition* of a specific mood or opinion.

As an example of how commented stage directions are used, the following virtual conversation, quoted in (26) (from Menges 1996), combines most of the strategies described in this heading:

(26) [Joyce_] everyone agree the list is good enough to vote on?
[lizzie] no
[Joyce_] please feel free to discuss further guys :)
* Joyce_ wait and watches with interest
[JmpMstr] voting to find the top six???
[dori2] lizzie describing actions is more than just making actions - it is a more detailed thing
[Joyce_] JM: yes, that's the next step
* OldBear wishes there were a chalk board or place we could write each item on yellow stickies and then cluster them into groups. This medium is not very good for developing true group hubris. ;)

[lizzie] dori, convince me
* JmpMstr lines up with OB… I agree
[dori2] well I can say dori2 sits
[lizzie] dori…. you mean… dori sits on the chair hapazardly
[dori2] that is simply using /me command
[dori2] yes - that is closer lizzie
[lizzie] ooooooooooooooo, this group is creating a mind meld…
we both knew you were sitting
* lizzie gets spooked
* Joyce_ smiles

It is also convenient to remark that some programs for chat room management, for example the ones analysed in Johnova (2004), include a command called "action" (or "think," depending on the program itself) that one can click on and the message that the user has typed is immediately turned into a commented stage direction. For example, if a user called <stunt> types "throws ice all over the place to freshen up his friends" and then clicks on "action," the program automatically yields the stage direction "*<stunt> throws ice all over the place to freshen up his friends*." In this way, the messages typed by users get mixed up with commented stage directions which, although they have also been typed by these users, exhibit a different format automatically created by the software, as in this example from Herring (2001):

(27) <Dobbs> come on, Danielle!!
<Danielle> No.
<Danielle> You have to SEDUCE me…
*** Action: jazzman reaches out for Danielle's soft hand.
*** Danielle has left channel #netsex
*** Action: Dobbs whispers sweet nothings in Danielle's ear
*** Action: Butthead moves closer to Danielle
<jazzman> danielle's gone dumbass

2. *Autonomous stage direction*. It occurs when the user's nonverbal behaviour is expressed with its closest translation, normally in one or two words, and framed by asterisks that separate it for the verbal content that they accompany. It is defined by Herring (forthcoming) as "predications that can function alone as complete performative utterances." An example is quoted below:

(28) <bull> What you're saying is funny *laugh*

In general, users resort to typical terms that one can find in a dictionary as prototypical of the equivalent nonverbal behaviour. Therefore, they vary inter-culturally depending on the lexical repertoire that languages offer for the description of nonverbal behaviour.

Again, these words between asterisks remind us of theatrical stage directions, and are typically used for saying hello and good-bye in chat rooms (Werry 1996: 60):

(29) a. <ariadnne> A N N E M A R I E!!!! *hugs*
<amya> *hugs* :)
b. <Untio> *kisses*
c. <Lola> I was joking *smiles*

2.5 Oralized written text

2.5.1 *Text-based chat rooms in the oral/written continuum*

The previous analysis has demonstrated that chat room discourse is a hybrid, somewhere between the stability (and often formality) of typed text, on the one hand, and the ephemeral (and often informal) quality of speech, on the other. As Baron (2009: 107) summarizes, although chat room discourse and other forms of computer-mediated text "are technically forms of writing, most varieties of online communication have often been thought of as forms of speech, with creative punctuation and typography substituting for paralinguistic cues (such as volume, proxemics, and facial expression) for expressing emotion." Therefore, many prototypical aspects of oral conversations can also be found in chat room interactions. Voiskounsky (1997) has summarized the oral and written qualities of computer-mediated communication (see Table 5.2) (see also English 1999, Kling 1996b and Kolko & Reid 1998: 213, 220).

Table 5.2 Oral and written aspects of computer-mediated communication

Oral	Written
Informal conversational style (e.g. use of first person, colloquial forms of address, recurrent idioms).	The user controls the composing process and the interlocutor cannot intervene in this process.
Search for textual equivalents to features of conversational interactions.	More complicated syntax than in oral conversations.
Short messages.	In the text, the "addresser user" makes the reason of the message explicit.
Intimate and emotional topic of conversation, which provokes the use of non-standard spellings and the use of symbols that are so typical of virtual conversations.	Possibility of revising the text and correcting errors before sending the message.

The anomalous situation of chat room discourse because of the simultaneity of oral and written features has led some authors to propose a new status for this discourse, a *third element* to be added to the traditional oral/written dichotomy, a hybrid that oscillates between the two extremes. In a similar fashion, I propose the label *oralized written text* for chat room discourse.[14]

Furthermore, Shank (1993) calls chat room discourse *multilogue*. As can be seen in Table 5.3 (Patterson 1996: Chapter 3), this *multilogue* differs from other types of interaction such as monologue, dialogue or discussion.

Table 5.3 *Multilogue* compared to other forms of virtual communication

	Monologue	**Dialogue**	**Discussion**	**Multilogue**
sender	one	one	initially one, maintains control	initially one, no control
receiver	one or more, passive	one, active	one or more, active	one or more, active
channel	FtF, mass media, other mediation	FtF or mediated	FtF or mediated	computer-mediated
examples	lecture, TV, books, radio, mailing list	FtF, letters, telephone, email	FtF, moderated groups	chat rooms, MUDS, news-groups, listserv

2.5.2 *Language games*

In headings 2.3 and 2.4 above I commented upon different strategies for the compensation of the loss of nonverbal features, vocal and visual, in text-based chat rooms. These strategies alter the typed text, and hence there is a presupposition that the interlocutors will choose the intended interpretation despite the increased effort that these strategies might posit. In text-based chat rooms, users make hypotheses about other users' ability to access a context in which the interpretation of the utterance will be relevant in the balance of cognitive effects and mental effort. Sometimes, the deviations from neutral text will demand extra processing effort, but this may be offset with additional interest (effects) from impressions (weak implicatures) such as the feeling of sharing the conventions of oralization, of an increase in sociability, or in humorous effects and playful atmospheres that are often generated in these text-based conversations.

---

14. Werry (1996) calls it *interactive written text* following Ferrara et al. (1991). In Elmer-Dewitt (1994) it is labelled *written speech*. Young (1994) calls it *writing conversation*. Merchant (2001) and Blanco Rodríguez (2002) prefer the term *written conversation*. Stein (2006) proposes *typewritten conversation*. Borreguero (2002, in López Quero 2010: 174) calls it *written simulations of oral conversations*. Finally, Fraca de Barrera (2007: 30) proposes the alternative term *pluridialogue*.

Among the most significant strategies of text deformation (and signals of oralized written text), the following can be listed:

1. *Orthographic mistakes*. Chat room discourse exhibits an informal style and, due to the pressure to type and send the messages as fast as possible, it abounds in orthographic mistakes. Some are involuntary, produced by errors in pressing the keys, but others are intentionally produced, as part of the language games that this medium favours (Mayans 2002b). Chat room users usually accept these mistakes as an inherent element of text-based interactions, although a user may occasionally complain if another user constantly makes too many mistakes.

2. *Phonetic orthography and eye dialect*. There are several strategies for the oralization of chat room messages. Following the terminological proposal by Androutsopoulos (2000: 521–522), some of them can also be found in chat rooms:

a. *Phonetic spellings*. It is the strategy of reproducing textually the text as it would be pronounced orally (for example, writing "imeil" instead of "e-mail" in Spanish). Chat rooms are prone to this strategy due to their hybrid oral-written quality.

b. *Colloquial spellings*. It refers to a colloquial reduction of words due to their pronunciation in the flow of speech. An example would be to transcribe strong and weak forms, as in (30):

(30) What d'yu wanna do?

c. *Regiolectal spellings*. These are transcriptions of regional variations of a language. It is also called *eye dialect*, an attempt to represent textually the phonetic qualities of a specific dialect, which involves phonetic elisions and text deformation. In (31a) there is an example to convey the typical accent of people living in the countryside, compared to a standard equivalent in (31b):

(31) a. You raaamblers don' realoyse the daaamage 'ee derz, traamplin' everywhoyre, leavin' gates open, with yourn dawgs runnin' oout o'control.
("Farmer Palmer," from *Viz*, quoted in Yus 1995: 56)

b. You ramblers don't realise the damage you do, trampling everywhere, leaving gates open, with your dogs running out of control.

d. *Prosodic spellings*. It refers to the textual transcription of prosodic contours of the voice by resorting to repetition of letters, capitalization, and the creative use of punctuation marks. This strategy is also frequent in instant messaging and *Twitter*, as can be observed from these examples taken from the corpus of *tweets* mentioned in Chapter 4:

(32) a. hasta la nocheeeeee (T4).
[*see you tonight*].

b. @usuario NOOOOO! Me rehuso a creerlo! (T5).
[*@username NO I can't believe it*].

c. Por fin juevesssssssssssss!!!!!!!!!!!!!!!! (T9).
[*Glad it's Thursday!*].

e. *Interlingual spellings*. This is a strategy that consists in transferring the phonetic attributes of a word from a foreign language but making it fit the orthographic conventions of the importing language. An example is the placement in Spanish of "e-" before several imported words from English (standard = estándar).

f. *Homophone spellings*. It refers to two kinds of textual alterations that do not correspond to parallel phonetic alterations that justify them:

f.1 *Lexical substitutions*. The strategy consists in writing a word (or part of a word) whose pronunciation is similar or equivalent to the initial word, but is shorter and complies better with speed-obsessed chat room users:

Word used	**Standard equivalent**
every1	everyone
18er	later
c u 18er	see you later

f.2 *Grapheme substitutions*. This strategy aims to replace one grapheme with another. In Spanish it is typical to replace "qu" with "k" (Mayans 2002a):

(33) <DUDU> jorrrrrr?
<DUDU> ke es esto? [*¿qué es esto? what is this?*]
<^XcyOnE^> es la cabeza de KaOs
<karin> diosssssssssss

It is quite clear that, for these chat room users, language is a flexible tool for transcribing speech, and they seem to obtain special pleasure from playing with the possibilities that the keyboard offers, a pleasure that certainly offsets any effort demanded in exchange.

3. *Abbreviations, acronyms, clippings*. Chat rooms are also full of these, turning paragraphs into a kind of hieroglyphic that only those users who master the conventions of these textual strategies can decipher. An example of text deformation is quoted in Belda Medina (2000: 562), which is now compared to its standard version:

Chat room conversation	Standard equivalent
– Hiya, r u new?	– Hello, are you new?
– Hi, no dewd	– Hi, no dude
– OK, a/s/l plz?	– OK, age/sex/location please?
– Lisa 20yo fm LA, u?	– Lisa 20 years old woman Los Angeles, you?
– Brad 22 m NY	– Brad 22 male New York
– kewl, gotta a selfpic for trade?	– Cool, have you got a photo of yourself to exchange?
– Yeap, sure, but what kinda format is urs?	– Yes, sure, but what kind of format is yours?
– Mine is a JPEG but I can't DCC, doesz't work. I can e-mail it though	– Mine is a JPEG but I can't send it via DCC, it doesn't work. I can e-mail it though
– Kewl, send it dewd, you first	– Cool, send it dude, you first
– Okay, g or x?	– Okay, general or x-rated?
– What? sorry, not into that stuff, bye	– What? sorry, not into that stuff, bye
– no prob, C U	– no problem, see you

The fact that only experts in these conventions are capable of understanding this conversation can be, in itself, utterly relevant to the users. The communicative success enlarges the portion of their environments that is mutual. Besides, chat rooms provide users with many impressions, in the shape of weak implicatures, that build up a particular source of satisfaction in these text-based interactions.

Abbreviations are also frequent in interactions. Again, some of them are raised to the status of conventions and enter the code of chat room discourse:

(34) msg (message)
thx (thanks)
tlk (talk)
bs (bullshit)

The problems of comprehension of *oralized written texts* increase with the use of acronyms that require *cyber-literacy*. These are usually integrated inside the messages, as in (35), re-written in (36) as standard English (Bunting 1999):

(35) BTW IMHO you deserve a :-* i'm sure you get plenty ITRW ;-). From my POV, you sound luscious (:-*;-*) and I'd like to spend time with you 24/7.

(36) By the way, in my humble opinion you deserve a kiss. I'm sure you get plenty in the real world. Flirtatious wink. From my point of view, you sound luscious (kiss, kiss) and I'd like to spend time with you, twenty four hours a day, seven days a week.

Another example is found in (37), published in Knight Ridder Newspapers (2002, quoted in Squires 2010: 467):

(37) RU der? GR8. Let's TLK bout all d abbrz & othr shrtcts poppin up mo&mo n MSGS… 'Are you there? great. let's talk about all the abbreviations and other shortcuts popping up more and more in messages. This might look like a word jumble -unless you're younger than 25, in which case you know it's the lingo used by kids to communicate with buddies.

4. *Ellipsis*. When users want to save time when typing their messages within such a dynamic environment, ellipsis is a useful tool. In everyday conversations, people usually leave implicit, non-coded, all the information that they expect that their interlocutors will be able to retrieve from context by themselves (an assumption of mutual manifestness). Needless to say, the more information is left implicit, the greater the role of the addressee and his/her responsibility in the eventual communicative success.

Ellipsis in chat rooms is used with a similar purpose, but time constraints must be added to the reasons for using it. Since chat room messages always contain the *nick* of the user at the beginning, the most frequently elided element is the first-person pronoun, more connoted in English than in Spanish because in the latter it is acceptable to omit this pronoun.

2.6 Attitudes and emotions in chat rooms

Oralized written texts can be a useful means to convey attitudes, feelings and emotions that are hard to code in neutral typed text. In this case, it is important to see to what extent this *text oralization* or *text deformation* (e.g. creative use of punctuation, capitalization and use of emoticons), as will be generically labelled in this chapter, allows "addressee users" to infer the intended interpretation of these attitudes and emotions correctly and also to measure their intensity (that is, to engage in what I call an *ad hoc measurement* of this intensity, see Yus 2005a).

Four hypotheses can be considered:

Hypothesis 1. Text deformation helps readers to identify the propositional attitude that underlies the composition of a message, that is, it has a *procedural*[15] role

15. The term *procedural* comes from the *conceptual / procedural* dichotomy. According to relevance theory, most words encode concepts and possess, therefore, a *conceptual* status. By contrast, some words such as connectives only code inferential "instructions" that aid the hearer in the relevant interpretation of the utterance, assuming a role of facilitator, a *procedural* role (W&S 1993, Blakemore 1987, 1992). For example, the connective "but" codes the instruction that the subsequent text should be interpreted as a contrast to the preceding one. See Yus (1998a: 328–329, 2010a) for a summary of this terminological dichotomy. See also Yus (2000a: thematic section 5.2) for a list of studies that have addressed this relevance-theoretic dichotomy.

facilitating the identification of this attitude. Besides, different amounts of text should be correlated to the intensity of these attitudes.

Hypothesis 2. Text deformation is useful to communicate propositional attitudes, especially when the content of the message is not sufficiently explicit to communicate them adequately. Again, it is also hypothesized that the more textual deformation there is, the higher the intensity of the attitude adopted by the chat room user.

Hypotheses 3 and 4. Users who want to communicate feelings (hypothesis 3) and emotions (hypothesis 4) resort to text deformation to assure a more accurate interpretation, and will engage in as much text deformation as the intensity of the feeling or emotion demands.

To test these hypotheses, 1.700 chat room messages were compiled, mainly from the Spanish chat room portal *Terra* (www.terra.es/chat). Besides, a questionnaire was handed out to students in order to corroborate or refute these hypotheses.[16]

2.6.1 *Hypothesis 1: Ad hoc measurement of procedural content*

This hypothesis tests whether text deformation works as a procedural element that guides readers in their inferential steps towards the interpretation of the subsequent part of the message that follows this deformed text. By doing that, the "addresser users" make sure that their attitudes are correctly identified and, at the same time, in an effort-relieving way in terms of mental effort. Kneepkens & Zwaan (1994: 129) add that "the emotional impression directs the attention of readers and helps them to decide which information is relevant for the situation and must be activated. This role of emotions is especially important when there are few textual and contextual cues, for example, at the beginning of a text." If textual deformation manages to convey a more precise account of the users' underlying attitudes in communicating the adjacent stretch of discourse, then the procedural role of some verbal elements will probably be enhanced.

From the analysis of the corpus, it can be concluded that most users resort to interjections for this procedural role of text deformation, in a similar way to their use in face-to-face interactions. Indeed, Wharton (2000: 194, 2009) concluded

---

16. The students lived in Alicante (Spain), studied at a High School (58.6%) or at university (41.4%). Most of them knew how to surf the Net and use the computer (77.52%) and how to use the chat room software and engage in text-based interactions (19.52% of them entered chat rooms on a daily basis). The students' ages ranged from 14 to over 21: 14–15 years old (34.32%), 16–17 (24.26%), 18–19 (4.73%), 20–21 (15.97%), and over 21 (20.72%).

that interjections can guide the addressee procedurally in the identification of the attitude regarding the subsequent text and activate several attitudinal concepts. From this point of view, saying "*wow!*" does not code a concept that the hearer interprets as "X is delighted"; rather, "*wow!*" activates a number of attitudinal descriptions that include, for example, happiness, surprise, or excitement. Something similar happens with the interjections found in the corpus of chat room messages, as in these examples:

(38) a. <stefany> joooooooo pronto empieza el curso.
[*blimey! school starts soon*].

b. <Jun-> aggg el brecol ta mu malo [i.e. el brécol está muy malo].
[*aggg broccoli tastes awful*].

In (38), "joooooooo" and "aggg" activate in the reader some attitudinal schemas, into which the text that is typed afterwards is inserted. The reader of (38a) is warned that he/she has to interpret the message under the schema "U regrets that *p,*" "U" being "user" and "*p*" being "school starts soon." Similarly, the reader of (38b) is warned that the correct attitudinal schema for interpretation is "U is disgusted by *p,*" "U" being again "user" and "*p*" being "broccoli."

Regarding the relationship between the amount of text typed and the intensity in the attitude, the hypothesis was that the informants would find differences between the neutral "jo" and "ag" and the text-deformation connoted interjections in (38). That is, it was predicted that the informants would interpret a supplementary layer of meaning relating intensity and amount of typed text, so that in (38) they would interpret a more intense "U regrets *a lot* that *p*" and "U is *utterly* disgusted by *p*," respectively. That hypothesis was confirmed in the questionnaire. The informants noticed differences between the neutral and the connoted versions, the latter revealing a more intense attitude. However, they found no differences between the version quoted in (38) (joooooooo / aggg) and a version with greater text deformation (jooooooooooo / agggggg). In other words, they identified a difference between the neutral and text-deformed versions but not between two text-deformed versions, as if text deformation worked as an "either-or" strategy: if text deformation is applied, a supplementary layer of intensity is added during interpretation, but no levels of intensity are paired to quantity of text deformation.

2.6.2 *Hypothesis 2: Ad hoc measurement of propositional attitude*

Users of text-based chat rooms (and hearers in face-to-face interactions) insert the proposition expressed by the utterance into a schema that indicates the user's (or speaker's) attitude underlying this utterance. There are several ways in which

this attitude can be communicated. These include that-clauses introduced by an attitudinal verb, as in (39a), parenthetical clauses (39b), verbal moods (39c), illocutionary adverbials (39d), and evidentials (39e), among others:

(39) a. I regret that you failed your exam.
b. It's time to go, I guess.
c. Come here right now!
d. Frankly, I am not surprised.
e. No doubt, he is the best candidate for the job.

Most of them are also available to chat room users, as can be observed in (40b–d) for the attitude described in (40a):

(40) a. [*User B is delighted with what user A has just typed*].
b. <B> :-D (an emoticon that replaces a smile).
c. <B> Fantastic!!!!! (writing about the attitude).
d. <B> Wow!!!!! (an interjection halfway between the mere expression of feelings and coded communication)

On other occasions, though, as pointed out in Yus (2005a: 151), the underlying attitude (and also the feelings or emotions attached to the utterance) can only be derived from non-linguistic evidence, by focussing on vocal aspects (e.g. intonation) and visual aspects (e.g. facial gestures) of the specific context in which the act of communication is taking place. For instance, Carston's (2002: 156) subsentential example "Out!" in (41a) can be interpreted as an order, that is, as having the higher-order explicatures in (41b–c), depending on non-linguistically-coded paralinguistic information (and also contextual information about power relations between the interlocutors):

(41) a. Out!
b. The speaker is telling the addressee to get out of the room.
c. It is desirable to the speaker that the addressee get out of the room.

However, in my opinion (41b–c) do not fully reflect the kind of information that the speaker really intends to communicate with (41a). The speaker is probably also interested in conveying his/her feeling(s) when uttering "Out!" (e.g. anger, indignation, irritation), which may be highly relevant for the interlocutors' background knowledge on the kind of relationship which they hold. In face-to-face contexts, the combination of vocal cues (e.g. shouting) and visual cues (e.g. facial gestures) is a good resource for speakers to convey the intended extent of their feelings and emotions while uttering a command, but in virtual environments such as text-based chat rooms, letters and punctuation marks are the only resources available for this task.

In a similar way, the chat room user's propositional attitude is sometimes not "visible" (i.e. coded) and text deformation is useful to compensate for this non-coded information. In the corpus, this use of text deformation is frequently found:

(42) a. <Jun-> no chillesssssssssssssssss.
[*don't shout!*].

b. <ZePeLiNa> NO XHILLE OTIAAAAAAAAAA [¡no chilles, hostia!].
[*don't shout, damn it!*].

c. <^LoBe^> maaaaaan echao!!!
[*I have been kicked out (of the chat room)*].

d. <elia> malhe nooooooooo te vayas.
[*malhe don't go*].

e. <rey-swin> [Hazlo] pero yaaaaaaaaaaaaaaaa.
[*(do it) right now*].

Examples like these indicate that capital letters, repetition of letters and punctuation marks can aid in communicating propositional attitudes more accurately. These attitudes have to be interpreted and measured "ad hoc" during the fast exchanges of chat room conversations. A common example is the repetition of the punctuation mark, because it helps to recover both the user's attitude and its intensity. For example, given the first part of message (43a), (43b) would be the proposition obtained after inferential enrichment of its schematic logical form, whereas (43c) would be a higher-level explicature that provides an attitudinal schema favoured by the use of the punctuation mark. Finally, (43d) would be the proposition after *ad hoc measurement* of its intensity has been applied to the repetition of this punctuation mark:

(43) a. <beckham15msn> alguna valenciana??????????????????
[*Any girl from Valencia?*].

b. [*Is there*] any girl from Valencia [*connected to this channel*] [*who would like to chat with me*]?

c. <beckham15msn> is asking if there is a girl from Valencia [*connected to this channel*] [*who would like to chat with him*].

d. <beckham15msn> is asking *with insistence* if there is a girl from Valencia [*connected to this channel*] [*who would like to chat with him*].

The results of the questionnaire corroborated that users did add a supplementary layer of intensity in the attitude associated with the repetition of the punctuation mark. However, they did not, again, relate the amount of typed text to the degree

of intensity, and contrary to my predictions.[17] In general, it can be concluded that users find neutral text insufficient to communicate attitudes and emotions and text deformation is a useful discursive tool to ensure their correct interpretation. For example, in (44a) the text itself reveals the higher-level explicature that includes an attitude of begging. On the contrary, (44b) might be interpreted as an informative text if its text deformation did not indicate a negative attitude towards being fat. Finally, in (44c) capitalization is used (although capitalization is not a polite strategy in chat rooms) to communicate how desperate for an answer the user is:

(44) a. <Quesalid> bastaaaaaa… Bastaaaa pofavooo [por favor].
[*stop, stop please*].

b. <saratogo> estoy gorrrrrrrrrddddooooooooooooooooooooooo [gordo].
[*I am fat*].

c. <rubita69> DEJAMEEEEEEEE DEJARMMEE VEROSSSSSSSSSS
[*let me see you*].

2.6.3 *Hypothesis 3: Ad hoc measurement of affective attitude (feelings)*
In this case, textual deformation is a consequence of the need to communicate feelings in a more accurate way, given the limited options of language for that task. Again, the prediction was that there would be a coupling between the intensity of the feeling and the amount of text typed. In the analysis of the corpus, it can be observed that some users simply type the utterance that best describes the feeling that they hold, as in (45a), while others use text deformation to connote the feeling with a greater intensity (e.g. 45b), or even use conventions such as "Z" to show boredom, as in (45c):

(45) a. <ferrari> me siento solo.
[*I feel lonely*].

b. <morena> me aburrooooooooooooooooooooooooooooo.
[*I'm bored*].

c. <solito_casa> zzzZZZzzzz.

17. An explanation for this may be that my informants were reading the messages. Several studies have concluded that, in general, readers can only infer general categories of feelings and emotions from the texts that they read, and are rather inefficient when it comes to identifying subtle variations of intensity in the same feeling or emotion (see Gygax et al. 2003).

Interjections are also a recurrent means to connote utterances with an affective layer, as in the feeling of surprise communicated in (46) as "U is surprised while typing *p*":

(46) a. <sigma> uy!! ya se a parado esto.
[*hey this (computer) has frozen*].

b. <Malhe> uys esta celosita esta cambiando las buenas costumbres.
[*hey this celosita is changing good habits*].

Other examples are quoted in (47), together with the interpretation of the underlying affective attitude ("U" for user and "M" for message), connoted with higher intensity thanks to the repetition of characters:

(47) a. <xica_gogo> abridme un privado por favor!!!!!!
[*Send me a private message, please*].
[*U is feeling very eager while typing M*].

b. <yo> HOLAAAAA ALGUIEN QUIEREE HABLARR CONNMIGOO QUEEEE YO ACEPTOOOOO A CUALQUIERAAAAA.
[*Hi! Does anybody want to talk with me? I'll accept anyone*].
[*U is feeling incredibly anxious while typing M*].

c. <cufi> t kieroooooooooooooooooooooooooooooooooo
[te quiero].
[*I love you*].
[*U is feeling absolutely in love while typing M*].

d. <BEFLY> vaya cuerdas vocaleeeeeeeeeeeeeeeeeeeeeeeeeeeee
[*What vocal chords!*].
[*U is feeling utterly amazed while typing M*].

The analysis of the questionnaire also revealed that users interpret greater intensity in the feeling when text deformation is used. Informants were also asked about a possible relationship between the amount of typed text and the intensity of the feeling, that is, whether in (48) the intensity of the feeling when typing "hola" (*hello*) was higher in <sevillana14> (48f) than in <patricia> (48c):

(48) a. <Elitrix> holaa
b. <RuBiOWaPo> HoLaa!!!!
c. <patricia> holaiiiiiiiiiii
d. <tere_rubia> ola a todo el mundooooooo.
[*hello everybody*].
e. <chico_20> hhhhhhoooooooooooollllllllllllllllllaaaaaaaaaaaa
f. <sevillana14> holaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

Again, and against my predictions, the informants found no difference in intensity regarding the amount of typed text. This corroborates the idea, already concluded after previous hypotheses, that users can only distinguish between neutral and textually deformed text. Either a supplementary level of intensity is communicated or no intensity is communicated at all.

2.6.4 *Hypothesis 4: Ad hoc measurement of emotions*
It is very likely that users will type text-deformed utterances in order to communicate their emotions more accurately and connote them with some intensity.

There has been an on-going debate on the differentiation of feelings, attitudes and emotions. On paper, attitudes seem to be cognitively more lasting than emotions, even though one can temporarily hold a certain attitude to what one is saying (propositional attitude) or feeling (affective attitude). As pointed out by Pilkington (2000: 152–153),

> an emotion, such as fear or anger, is a temporary state, a response to some perceived event or state of affairs in the world […] An attitude, such as love or hate, involves the storage of a belief and/or phenomenal state in long-term memory, attached to a conceptual address […] Whereas an emotion is a temporary response to a situation involving the creation of a new desire or the strengthening of an existing desire, an attitude is focussed upon a particular object.

In general, emotions can be defined as "acute, intense, and typically brief psychophysiological changes that result from a response to a meaningful situation in one's environment" (E. Rosenberg 1998, quoted in Kidron & Kuzar 2002: 130). Besides, it is interesting to stress that emotions can be divided into two broad categories: those that are produced unintentionally (as part of *emotional behaviour*) and those that are produced intentionally (belonging to *emotive communication*).[18] The former are usually inter-culturally valid, whereas the latter are influenced by the culture to which the individual belongs, which regulates their (in)adequate uses in interactions.

Generally, text-based chat room users can only communicate emotions intentionally (there is usually no unintended typing on the keyboard), but other users can infer emotions from text beyond the user's control. This fact introduces a danger of misinterpretation of emotions compared to the straightforward access to emotional cues in face-to-face interactions. As Riordan & Kreuz (2010: 167) point out,

---

18. Caffi & Janney (1994: 328) characterize these types as "a type of spontaneous, unintentional leakage or bursting out of emotion in speech" (*emotional*) and "the intentional strategic signaling of affective information in speech and writing in order to influence partners' interpretations of situations and reach different goals" (*emotive*).

> the ease of interpreting nonverbal cues in FTF [face-to-face] relative to CMC [computer-mediated communication] may be the result of socialization skills accumulated by FTF contact over a lifetime; as CMC is relatively new compared to FTF socializing, it is still a channel in which cue use and interpretation is negotiated between users and learning to encode and interpret emotions using these cues is an ongoing experience.

But emotions abound in this virtual environment where anonymity and lack of physical co-presence allow for emotional exaltation without the control that society usually exerts on individuals in physical settings, something that also applies to social networking sites (see Blincoe 2009, Derks et al. 2008). The analysis of the corpus reveals that text deformation is systematically used for communicating emotions and, perhaps, also their intensity. The most typical emotions are the easiest to type, those that often *alternate* with speech (Poyatos 1975), such as laughter (49) and shouting (50):

(49) <xicagogo16> os kiero muchisimo!! jeje
[*I love you a lot!! ha ha*].
<ZePeLiNa> Quesaild: jajajajajajajja jajajaajajajajjajajaja.

(50) <cristina_c_a> aaaaaaaaaaah.
<ESIGUAL> waaaaaaaaaaaaaaaaa.
<zuMBaooo> WweeeeeeeeeeeeeeeEEEEEEEEEEEEEeeeeeeeeeee.

Emoticons are also useful for conveying emotions in chat rooms, especially broad categories such as happiness [:-)] and sadness [:-( ].[19] It would be useful, in this sense, to test whether the repetition of an element of the emoticon communicates a higher intensity in the emotion or not. One of the questions in the questionnaire dealt with this hypothesis. The informants were asked whether they found differences in emotional intensity between (51a) and (51b) (prototypical emoticon vs. connoted emoticon with repetition of one of its elements). Then, they were asked whether they also found differences between (51b) and (51c) (intensity of the emotion related to the amount of typed characters):

---

19. Although there is a large variety of emoticons, only a handful of them are frequently used by Internet users, especially the ones communicating happiness [:-)], wink [;-)] and sadness [:-( ] (see Walther & D'Addario 2001). For instance, Schulze's (1999) quantitative analysis revealed that only nine emoticons are extensively used. Taking into account the process of stabilization that certain emoticons have undergone, especially with the online publication of various emoticon glossaries, it can be stated that any repetition of one or more of the typographic signs composing the emoticon should be taken as a connotation of the default type, and hence the other users reading these emoticons will tend to infer an additional layer of meaning associated with this repetition.

(51) a. <Quesalid> :-)
b. <Quesalid> :-)))))
c. <Quesalid> :-))))))))))))

The answers yielded results that are consistent with the informants' answers for previous hypotheses. Again, most of the users found high intensity in the emotion communicated by (51b) compared to the neutral (51a), but found no differences between (51b) and (51c). This is consistent with previous conclusions that corroborate that identifying attitudes, feelings and emotions in chat rooms is an "either-or inference," in the sense that either they are interpreted as neutral (no text deformation) or a supplementary level of intensity is inferred (via text deformation), but beyond this initial dichotomy, there are no fine-grained inferences of intensity related to parallel amounts of typed text.

3. See you on messenger

The title of this heading has turned into a typical phrase among adolescents, who have replaced the physical "see you" (in a bar, square, or park) with an electronic equivalent within *instant messaging* (henceforth IM). Ours is a highly technified society with a convergence (and clash) of increasingly virtualized physical settings for interactions and a growing *physicalization* of virtual environments for interactions. Since communication has evolved towards a total hybridization of physical-virtual scenarios in personal networks (Yus 2007b), IM allows for synchronic conversations that substitute physical interactions or complement them efficiently (see Baron 2008c).

3.1 Instant messaging compared to other forms of interaction on the Net

IM shares attributes with other forms of synchronous virtual interactions. Specifically, it shares with chat rooms many characteristics such as the fact that they are still typically text-based but have the possibility of visual and audio contextualization (web cam, microphone) (see Peter et al. 2007). Besides, both of them generate lists of messages in a central area of the screen while other private conversations are taking place in multiple windows, and they both exhibit similar strategies of oralization and text deformation. However, there are also differences between chat rooms and IM:

1. Chat rooms typically hold one-to-many interactions, even though private conversations can be held in multiple windows. IM is typically a one-to-one interaction, as can be seen in Table 5.4 (adapted from Baron 2008c). But, again, it

is also possible to engage in multiple conversational threads in the same area of the screen.

Table 5.4 Instant messaging compared to other forms of Internet-mediated communication

	Synchronous	**Asynchronous**
One to one	Instant messaging	E-mail, SMS
One to many	Chat rooms, videoconferencing, blogs	Newsgroups, e-mail distribution lists

2. On most occasions, chat room users simply access a website that contains the interface (for example http://terra.es/chat/). The user can have the window of the chat room minimized without interfering with the user's task at hand (although modern chat rooms produce noises such as telephone calls that do interfere). On the other hand, IM generates pop-up windows on the taskbar that warn users of an in-coming message while, at the same time, a sound is heard. From a pragmatic point of view, this "visual warning" is useful to stress the user's communicative intention underlying his/her informative intention. However, it also produces disturbances in the user's task at hand.

3. Chat rooms are available as soon as the user accesses the website that contains the interface, whereas IM demands the installation of one of specific software programmes that, at least for the time being, are not mutually compatible, in the sense that all users who maintain an IM conversation have to use the same software (but this is not a problem since most users install the same popular IM software).

4. The IM software warns the user when someone from his/her *buddy list* logs onto the system. In one of its most popular versions, a small window opens on the taskbar with a photo of the user and a message communicating that this user has logged onto the IM, and a sound is heard. By contrast, chat rooms do not offer information about contacts, but a text is offered in the main area of the chat room if a user with a *nick* has entered the room.

5. Chat rooms arouse users' interest mainly due to the possibility of meeting strangers that hide behind the *nick* (although it is also used for conversations with friends and acquaintances). IM is typically useful for dialogues with people that users already know and meet in physical settings on a regular basis (for example, users who go to the same college, high school or university, see Quan-Haase 2008). IM is therefore often a complement to face-to-face interactions. As a consequence, although IM interlocutors do use *nicks*, these do not differ substantially from their real names (see Dietrich 2004, Quan-Haase & Collins 2008). As Abril (2006: 39) qualifies, "unlike chat rooms and newsgroups, which

are macro-communities open to anybody who wants to leave a message, with IM each individual chooses how, when and with whom." A female user corroborates this feature when she comments that IM "is like being in a coffee shop to which friends and acquaintances enter to chat for a while […] and you avoid all the weird people that inhabit chat rooms" (in Abril ibid.).

6. Chat rooms are open to everybody, whereas only IM users who belong to the list of contacts (buddy list) can interact with a specific user.

7. IM offers options for telling other users about the situation of a user, especially the reason for a temporary absence (Cameron & Webster 2005: 86), often with connotations of group membership and personal status within a network of friends and acquaintances.

8. Chat rooms and IM share the users' strategy of text deformation and emoticon use. However, they are more frequent in IM, where conversations are richly decorated with a visual display of coloured emoticons and text.

9. As far as the use of distinctive avatars[20] is concerned, Kang & Yang (2006: 1177) state that IM avatars are more directly related to the user's personality and looks, while the ones used in chat rooms are more imaginative and disconnected from the user. Since IM users meet face-to-face in physical settings, they do not mind using avatars that resemble their physical appearance. By contrast, chat rooms are places for strangers and it comes as no surprise that avatars are so different from their actual users.

IM has also been compared to e-mail in spite of the fact that IM is synchronous and e-mail is asynchronous. Besides, the latter has a more "formal" connotation for adolescents, who use it only for contacting teachers and parents, whereas IM is more informal and relaxed (Grinter & Palen 2002). But despite the obvious differences, there is a similarity between them in terms of how long users have to wait for a reply (i.e. the *chronemics* of IM communication). Indeed, although there can be intense IM interactions composed of chained messages and multiple conversational threads, part of the users' knowledge of IM conventions includes the expectation that the interlocutors may take some time to answer a message, just as in e-mail communication (Nardi et al. 2000). Nevertheless, within IM interactions misunderstandings are frequent due to the differing conceptualizations of how long it is advisable and acceptable to wait until one's interlocutor replies, with negative effects on the regulation of IM turn-taking (Voida et al. 2002: 191).

Finally, IM has been compared to SMS *texting* (Ling & Baron 2007, Baron forthcoming). In principle, they seem to be difficult to compare, because IM is

---

20. Normally bi-dimensional avatars that cannot be altered. Later in this chapter I will differentiate between these bi-dimensional unchangeable avatars, which I will call *graphic avatars*, and 3D avatars that can move, express feelings, etc., labelled *corporeal avatars*.

normally performed between people located at specific places, in a synchronous way and typed from computer keyboards (although there are increasingly popular IM applications for smart phones), and *texting* is often performed by people located anywhere, asynchronously and typed from the small mobile phone keyboard. From a linguistic point of view, SMS messages tend to be longer than their IM counterparts, probably due to the habit of IM users to fragment messages into several micro-texts so as to maintain their "presence" in the IM interactive environment (more on this below). Besides, SMS messages usually contain more abbreviations due to the 160-character limit that can be typed in one single message.

3.2 Why use instant messaging?

IM is widely used among adolescents as a "natural" form of interaction on the Net and its use decreases at an older age. Therefore, it seems pertinent to determine the benefit IM offers adolescents to justify the massive use of this *cyber-medium*. Within the framework of this book, it is important to determine what expectations of relevance are satisfied with the use of IM. And both the expectations and the use of IM are influenced by the constant evolution of *cyber-media*, which affects not only the kind of use that is made of this medium, but also the assessment and resulting relevance of IM exchanges in ordinary situations. In this sense, Baron (forthcoming) summarizes some of the current characteristics of IM that can be considered evolutions from its initial design and applicability: (1) many IM conversations are conducted asynchronously (rendering them not "instant"), (2) IM now supports multi-person chats, (3) users can now be logged on to IM but "lurk," rendering them invisible to members of their buddy list, (4) some IM software offers voice and video options besides typed messaging, (5) IM is incorporated into other computer platforms, especially commercial web sites offering clients opportunities for live "chatting" with customer-service representatives, and (6) IM can now be done on mobile phones.[21]

In my opinion, IM offers adolescents a wide range of relevance-generating attributes, several options for personal reward both from an individual point of view (achievement of communicative purposes) and from a collective point of view (feelings of group or network membership, and satisfaction at sharing a virtual setting). But several reasons can also be found beyond the age constraint. For example, the IM interface is a user-friendly environment that offers immediate

21. IM is also one of the synchronous options for avatar-mediated conversations in 3D virtual worlds (see 4 below). It has also been incorporated to SNSs such as *Facebook* and *Tuenti*. Although SNS software labels this service as "chat," in reality it is an IM application for one-to-one conversations, even if many of them can be carried out in parallel.

reward for users who seek synchronous interactions on the Internet, and the effort associated with using the interface (which in previous chapters of the book has been signalled as a potential source of alterations of relevance) is reduced significantly (Chung & Nam 2007: 227). According to Grinter & Palen (2002) users' needs for interaction can be summarized as follows: (a) *needs for socializing* (informal conversations to spend some time with friends, with no pre-arranged topic), (b) *for event planning* (to arrange meetings, for example to go to the cinema), and (c) *for schoolwork collaboration* (to clarify problems with homework, for instance). Quan-Haase (2008: 109) adds a fourth need: to be able to engage in multiple one-to-one conversations simultaneously, an aspect that I prefer to address under the individual / group interface in the next heading.

3.3 The individual versus the group

Cognitive pragmatics is especially interested in IM as a tool for interactions that satisfies communicative needs and allows the possibility to draw relevant conclusions through inferential strategies that turn the schematic logical forms of messages into fully contextualized and meaningful interpretations. These interpretations may be relevant both in a purely informative sense and also as a source of phatic socialization, together with the reinforcement of group membership (see Boneva et al. 2006, Flanagin 2005). The individual / group dichotomy that IM fosters is one of the key elements that, in my opinion, explains why IM is massively used by adolescents and youngsters as part of their process of individual and social identity-shaping, and is not so widely used by adults. Another reason might be that the benefit that adolescents obtain from IM offsets the drawbacks involved in using this *cyber-medium*. By contrast, adults find it too annoying and distracting as it interferes with other tasks (Birnholtz 2010).

From an individual point of view, IM is a tool for fast synchronous communication that satisfies specific expectations of relevance with little mental effort in exchange, and with greater emphasis on interactions between users who already know one another in physical settings.[22] Besides, the content exchanged through IM, about apparently irrelevant topics, favours phatic strategies. These "useless" topics have, nevertheless, an impact on users' identities and social awareness. Therefore, if we establish a scale of communicative intentions, we will find that on IM very few messages are intended to satisfy individual needs; most of them

22. This function of "complement" that IM plays with respect to physical interactions makes it difficult to draw a clear dividing line between physical and virtual IM interactions, and it is also difficult to conclude which kind of interaction is more important to users in terms of intensity and impact (see Bryant et al. 2006: 586).

possess a connotation of satisfaction of collective communicative[23] intentions (a sort of *we-intention*), whose fulfilment demands the participation and cooperation of all the users who are synchronously logged onto the IM system at a specific moment.

From a group or social point of view, we can find many instances of IM interactions of a phatic kind, filled with (apparently) irrelevant utterances in a purely informative sense. But they do provide relevance in making mutually manifest assumptions such as awareness of co-presence inside the group or network of friends who are synchronously inter-connected, as well as relevance in the mutual manifestness of being present in the conversation, even if not actively participating. In IM conversations among young users, there is an obsession with demonstrating that the user is part of the interaction, part of the synchronous collectivity. A female user's comment (in Lewis & Fabos 2005: 487) is illustrative of this: "if I don't get on, like if it's broken, like if the Internet's not working, I'll, I'm like ahhhh! So I'll call my friends, and I'll be like "Who's on? What are you talking about?!… I'll be like, Get on my name and pretend you're me for a little bit!" This obsession with group recognition partly explains the habit of sending utterances as several chunks of text in successive messages (Baron 2010a). With this strategy, the user maintains other users' attention and awareness of his/her presence in the on-going interaction. This is why Baron (2010b) proposes a specific terminology that I think is worth quoting and in which there is no obligatory pairing of the utterance and the typed message:

1. *Transmission Unit*. An instant message that has been sent:

   (52) Tom: how are you doing, mate?

2. *Utterance*. A sentence or sentence fragment in IM:

   (53) a. Susan: I've just returned from a restaurant! [sentence].
   b. John: Come to think of it… [sentence fragment].

3. *Sequence*. One or more IM transmissions sent sequentially by the same person:

   (54) Mark: hi man!
   Mark: what are you up to?
   [this sequence equals two IM transmission units].

---

23. From a philosophical point of view, collective intentions are considered intentions to participate in a group to carry out an activity in which the participants feel that they are members of the group (Cheung et al. 2007). IM clearly fits this kind of satisfaction of collective intentions from the involvement of the users in the positive and effective development of virtual interactions.

4. *Closing*. A series of transmissions (between IM partners) at the end of an IM conversation, beginning with one party initiating closure and ending with termination of the IM connection:

(55) Sam: Hey! gotta go [first symptom that Sam is willing to terminate the conversation].
[…] [subsequent exchanged messages].
Sam: I'm off to work! [last transmission in the conversation].

5. *Utterance chunking*. Breaking a single IM utterance ("sentence") into two or more transmissions:

(56) Joan: that must feel nice
Joan: to be in love
Joan: in Spring

6. *Utterance break pair*. Two sequential transmissions that are grammatically part of the same utterance:

(57) Ally: what are you bringing to the dorm party
Ally: on Saturday?

It should be noted that several aspects of IM that have been addressed as mutually independent share, in reality, some attributes concerning their social connotation. I am referring to (1) *multi-tasking*, which is more frequent in IM than in other forms of Internet-mediated interactions; (2) *multi-windows*; (3) *buddy list management*; and (4) strategic use of automatic *away messages*, which are often personalized. I shall briefly comment on them below.

1. *Multi-tasking* has already been mentioned in this book. It refers to the simultaneous engagement in different activities, normally with parallel computer applications. It is typical of adolescents and very frequent in IM interactions, since young users frequently have several windows open for one-to-one IM conversations, plus the Internet browser and the word processor for doing homework.[24] An explanation of this high frequency of IM-related multi-tasking may

24. Nowadays there seems to be a shift in multi-tasking. Instead of using different programs for parallel tasks (e.g. word processor, Internet browser, IM software…), the trend now is to use the same portal for all kinds of multi-tasking activities on the Net. For example, as commented in Aldama (2011:6), in the Chinese social networking site *QQ*, which competes with *Facebook* in the number of registered users, people can use e-mail (*QQMail*) and a virtual hard disk (*Wangluo Yingpan*), update a blog (*QQZone*) or a microblog (*Tencent Weibo*), download music and ring tones (*QQYinyue*), buy goods online (*Paipai*) and play online games (*QQYouxi*). Users can also buy plane tickets, look for romance (*QQTongchang*) and even look after a virtual pet (*QQChoungwu*).

lie in adolescents' need of mutual manifestness of virtual co-presence of friends during an IM session (a sort of *ambient awareness*, as Thompson 2008 would call it). This need of mutuality prevents users from leaving the IM session, and forces them to reply to messages even though they are engaged in a different task (these messages are, predictably, short and schematic, worthwhile only for maintaining mutuality of co-presence).

Multi-tasking (and parallel multi-windows) may retard the user's replies to IM messages, but this time a gap between transmission and reply is accepted by the users, who do not expect immediate feedback and assume that the other user is probably doing something else while logged onto an IM session. This is what Baron (2008c) calls *language under the radar*. In other words, instead of devoting most of our cognitive resources to following and participating in IM interactions (as face-to-face conversations demand), IM is becoming something that is followed as a background to other activities within multi-tasking, just as background music might be playing while we are engaged in a different task, and the IM "volume" can be lowered or increased depending on how involved the user is in the current IM session. This idea was corroborated in another piece of research by Baron (forthcoming), where 98% of 158 informants (half male, half female) were busy doing at least one different task (on the computer or away from it) while participating in an IM interaction.

2. *Multi-windows* is clearly related to multi-tasking. It takes place when the user has opened several windows for one-to-one IM conversations, and the user tries to monitor and follow in a relevant manner several conversational threads with different people and about different topics simultaneously, despite the effort-producing challenge that this involves. This challenging activity refers again to the social need to make it clear to other users (i.e. to obtain a mutual manifestness) that the user is connected and able to sustain several interactions as a signal of sociability and as a source of prestige for other users. IM can indeed be a source of social positioning within the network of friends and acquaintances. Schwarz (2011:77) quotes the following illustrative example:

> A 16-year-old blogger published an IM conversation with a boy she didn't know who tried to lead the conversation to sex, while she played dumb, wittily using puns and double-meanings she found in his formulations to respond to all of his questions with innocent, non-sexual answers. Sharing the evidence of her victory, the girl won her friends' esteem.

A third of the informants in Boneva et al. (2006) commented that one of the strong features of IM is, precisely, the possibility of carrying on multiple conversations simultaneously, what Garrett & Danziger (2007) call *polychronic communication*. This "strong feature" certainly demands extra cognitive resources and

supplementary mental effort to maintain "interactive congruency" throughout all the conversations in these windows. It comes as no surprise that the adolescents interviewed in Grinter & Palen (2002) admitted that there is a limit to the number of opened windows beyond which it is impossible to maintain congruency or coherence. As a general rule, the average number of windows opened simultaneously is four, which users arrange in parallel on the computer screen as a mosaic of windows. The most typical strategy to maintain congruency and coherence in these multiple conversations is to follow with greater interest one or two of them and simply follow "in the background" the other conversations (see Lewis & Fabos 2005: 218–219 for an example).[25]

Hence, many of the IM exchanges are basically to say hello and carry a phatic connotation, i.e. they are short messages that stress a social connection and copresence in the virtual environment of IM, and they are useful for social grooming beyond their informativeness. In this sense, the other users assume that the user is surely participating in several conversations simultaneously and that they will have to fight for his/her attention by offering a presumption of relevant inferential outcomes.

3. The *buddy list* is the exponent of the contacts on the screen that belong to the user's personal networks. Therefore it is vitally important for the adolescent user. These friends and acquaintances are grouped under categories that reflect, with greater or lesser accuracy, the mental picture that the user has of his/her personal networks.

4. Finally, *away messages* are often personalized by users, and they regulate the degree of social involvement that they desire within IM interactions, even when the user is not logged onto the system. It is a tool for the management of interactive availability in a type of interaction where, as Bays (2010: 43) puts it, "everything from the onscreen activity with its colours and mosaic of windows to the physical environment of the user who may be listening to music, talking on the phone or engaged in exogenous conversations are also important to forming the whole IM experience." When the user personalizes messages, very often they aim at relevance with them, trying to call other users' attention, for example with humorous quotes (Nastri et al. 2006: 1027), or explaining in detail the reasons for

25. Emoticons can also be used in isolation for the purpose of manifesting presence in other users' windows while focussing more on a particular window. As Bays (2010: 57) points out, emoticons can "show continual presence in front of the screen and in a particular conversation, the participant can simply send a smiley as his turn, as a kind of conversation filler. This accords him time to scroll back up the conversation window to see what the topic was and how it evolved while he was away in order to answer the adjacent pair or contribute a relevant message to the topic. The ambiguity of its semantic meaning allows the smiley to be relevant in many situations and to retain the general tone of the conversation."

being absent. For example, one of the away messages quoted in Grinter & Palen (2002) reads "I'm currently removing all dirt, grime and other dead biological matter from my body. I can be found in the nearest decontamination center."

3.4 Oralized written text in instant messaging

The texts typed on IM exhibit similar strategies of oralization and text deformation that have been studied in the section devoted to chat rooms (see Varnhagen et al. 2009, Baron 2010a, 2010b, Bays 2010, Herring forthcoming). The evolution of chat rooms into a more contextualized *cyber-medium* incorporating web cam and sound is also valid for IM. The messages exchanged through IM exhibit oral properties, with short utterances and dynamic interactions, typical of ephemeral oral dialogues. But users also value the typical properties of written communication such as the lack of a need for immediate feedback. The informants' opinions quoted in Voida et al. (2002: 188) are enlightening, because they describe IM as

> being nearly synchronous but able to be attended to when opportune. The former characteristic is shared with most verbal communication; the latter, with most written communication. Implied in the interviews of our participants is that instant messaging is valued because of the unique balance it holds in affordances between the conventions of verbal and written communication.

Curiously, users also make typical gestures of face-to-face interactions while typing their IM messages, even though their interlocutors cannot see them most of the time (see Marcoccia et al. 2008). Bays (2010: 46) draws similar conclusions from the analysis of IM users' behaviour:

> When the new message appears, there is an immediate often physical reaction. Among others, we found laughs, shaking heads, pointing fingers at the screen, having a dumbfounded expression and moving lips to read the new message. Generally, these are the same physical attitudes that can be found in other conversational settings from face-to-face interaction to talking on the telephone, whereas this behaviour is rare (at most) for email or reading a blog.

Emoticons seem to be more abundant in IM than in chat rooms. Xu et al. (2007) have made an exhaustive analysis of emoticon use in IM. Initially, they proposed three basic uses: to accentuate or emphasize the meaning of a message, to convey the user's mood or impressions, and to enrich a verbal utterance with visual information. They then concluded that the greater or lesser frequency of emoticons in IM seems to be influenced by three factors: the degree of (in)formality of the conversation, the relationship that both interlocutors hold, and the personality of the interlocutor. These are commented upon below.

1. The degree of formality is usually established according to two main types of interaction: *task-oriented* and *socio-emotional*. In the former there is a clear objective and the exchange of information is prominent. In the latter, by contrast, the interaction is informal, of a phatic quality, with no predetermined topic, and dominated by the expression of feelings and emotions. It is not surprising that using emoticons in task-oriented communication is not considered appropriate and is seen as a source of unnecessary distraction, whereas in socio-emotional conversations their use is not only predictable, but expected, as happens in most informal conversations among adolescents.

2. There are several levels of relationships between users within IM. In the case of intimate friends, it is more likely that the users will be willing to express feelings and emotions while typing their messages, and emoticons are one of the resources at hand to communicate them more efficiently. The opposite occurs when the conversation takes place between acquaintances.

3. As far as the user's personality is concerned, the use of emoticons is acceptable or inappropriate depending on the interlocutor's personality, since some users do not like finding emoticons in the messages that they read.

4. Chatting in 3D: Advances, avatars and *Second Life*

Chat rooms and IM have evolved enormously in the last few years, even though text-based interactions are still very frequent. In fact, the chat room and IM interfaces have not changed dramatically over the years. In chat room interfaces there is still a wide central area for open messages, a list of *nicks* that one can click on for private conversations and a space for typing one's own messages. When typing, the interface now gives users options for fonts such as bold, italics and colour, and also pre-designed 3D or animated emoticons that were not available in the 90s, when chat rooms first became popular. Furthermore, several authors have proposed interfaces for a better contextualization of conversations (see Yus 2003f). Some of them are briefly reviewed below.

1. Viegas and Donath (Donath 1996, Viegas & Donath 1999, Donath et al. 1999) proposed an interface for chat rooms called *Chat Circles*, in which there is only one room for interactions where several circles represent the users who are logged onto the system. The user can move the circle on the screen, and these circles change size to accommodate the text typed inside. There is also a possibility of engaging in private conversations, since users can only chat with other users whose circles are close enough to theirs.

2. *BodyChat* (Vilhjálmsson & Cassell 1998) is another proposal of an interface. It uses anthropomorphic figures, known as *avatars*, to construct a more

contextualized virtual conversation (avatars are essential in interactions within virtual worlds such as *Second Life*, see below). Among the options for expressiveness, these figures can visually manifest that they are paying attention and also blink their eyes. The problem is that the design of avatars is very rudimentary; these avatars cannot reproduce even a minimum of gestures to the user's satisfaction, and hence they distract rather than aid in virtual conversations.

3. *Chatscape* (Lee 2001) offers a low-quality graphical environment that attempts to provide more contextualization in virtual interactions. Users' messages appear on the screen as comic-style balloons that arise from polygons whose shape can be varied in certain ways.

4. *The Palace* is a 2D chat room where users have to choose the bi-dimensional avatar (which I will call *graphic avatar* below) that best suits their personality. After making a choice, the user enters a room where text-based conversations are taking place. There are some options for personalization of the avatar. Besides, users can make their avatars "jump" around the room and get closer to other avatars, even though closeness to an avatar is no pre-condition for starting a conversation. The biggest limitation of this chat room is the impossibility of conveying nonverbal information with the fixed avatar.

Besides *The Palace*, other graphic chat rooms offer similar options for image-supported text-based conversations: (a) *The Manor* (www.madwolfsw.com/) offers several advances, for example the possibility of changing clothes and the availability of 15 animations. (b) *TowerChat* (www.towerchat.com/) is also a 2D environment, but with an original "bird's-eye view" over the scene, giving users the feeling of a 3D scenario. It is divided into eight spaces or "towers" (politics, music, love, and others). One can choose a male or a female avatar. (c) *Voodoo Chat* (www.voodoochat.com/) is similar to *The Palace* but requires the installation of special software on the computer. (d) *VPchat* (www.vpchat.com/) is similar to *Voodoo Chat*. It contains rooms with avatars that interact with one another through text, gesture and voice. (e) Humphrey (2009) proposes a version of chat room mediated by what she calls "masks," that is, avatars that users create in order to interact without the danger of visual exposure and also to play with a multiplicity of online identities. In her study, she concludes that through these masks some aspects of the users' personalities are inevitably distinguishable. (f) *V-Chat* (Smith, Farnham & Drucker 2000) is a program in which rooms have a 3D appearance and users (up to 25 simultaneously) type their messages. As in *The Palace*, users choose a 2D avatar. These may be chosen from a gallery of pre-determined avatars or they can be created by the users themselves. There are seven possible nonverbal behaviours: anger, sexual innuendo, sadness, gesture of puzzlement, silly face, smile and wave. All participants can read what other users have typed, but there is an option that allows a user to send personal "whispers."

5. *Comic Chat* (Kurlander et al. 1996), as its name indicates, is a program that automatically generates panels that resemble the ones in comic books, and inside which the users are visually represented, as well as their messages in the form of comic-style balloons (the users have previously selected a character). The problem with this kind of chat room is that the characters suffer from a lack of expressiveness, just as in most comics.

6. *Coterie* (Spiegel 2001) is limited from a contextual point of view. It is basically an accumulation of oval figures that represent the users who are logged onto the system and interacting. Their colour and proximity to other figures indicate visually who is interacting with whom and with what intensity.

7. Smith, Cadiz & Burkhalter (2000) propose an interface called *Threaded Chats*. It is a visual arrangement of "trees" of conversational threads that help users identify and follow them. The software automatically links a message to the one it refers to and inside the overall thread of the conversation.

8. Ryu (2008) proposes the integration, in the same IM window, of both the text that the user is typing and a three-dimensional avatar with the ability to convey a range of emotions. The author starts with an acknowledgment that users get too distracted if the avatar is fully animated, and that distraction worsens if the avatar and the text share the same space in the window. Therefore, the avatar should be as "non-intrusive" as possible. But, on the other hand, placing the avatar outside the IM window would entail losing some of the user's attention towards this avatar. This is why the avatar appears as a faint image and the user's text appears superimposed on this image. The avatar can communicate a number of emotions, which can be personalized to a certain extent.

9. Fabri et al. (2005) propose a software for IM interactions called *Virtual Messenger*. It adds an animated face to the typical window for typing text. This face is capable of communicating several emotions (it contains animations for the eyes, the eyebrows, the cheeks, the mouth, and the whole head). As a result, six universal facial expressions can be conveyed: joy, surprise, anger, fear, sadness and worry. One of the conclusions when testing this interface was that the "addressee users" tended to imitate the gestures produced by the animated face, a kind of *avatar empathy* that reproduces a typical quality of human cognition, since there is a biological and cognitive component in imitation, to the extent that we learn to behave in the world by imitating what others do. And we possess a number of "mirror neurons" that have evolved specifically to check what others are doing and that are connected directly to other areas of the brain in charge of movement and comprehension of the outside world. One limitation of this interface was that users have to produce nonverbal behaviour intentionally, whereas daily face-to-face interactions are full of *exuded* information that "leaks" from the person beyond his/her conscious control.

But undoubtedly, the most impressive evolution in virtual conversations is the development of 3D virtual worlds such as *Second Life* (www.secondlife.com, henceforth SL), which possesses interesting attributes for a pragmatic analysis of Internet-mediated communication.[26]

SL is a 3D virtual world where users, by means of three-dimensional *alter egos* or *avatars*, interact with other users-avatars. It is much more than a simple environment for interactions, though. In SL many parameters of physical scenarios are reproduced, including bank transfers, and the purchase of land, buildings, or clothes, using virtual money (later turned into real money). Within SL, it is the creative, interactive and social activity of users that qualifies SL as a "world" where many real-life parameters are reproduced.[27] Therefore, these users are *prosumers*, devoting as much time to consuming information and virtual goods as to producing them (D. E. Jones 2005). Some examples of how blurred the distinction between physical and virtual lives has become in these virtual worlds are quoted below:

> A 17-year-old Dutch teenager was arrested this week on suspicion of stealing furniture worth £2,800 from a hotel room. Four other teenagers were also questioned about the offence. It is believed they moved the stolen furniture into their own hotel rooms. Such a minor incident might not have merited a paragraph in the local paper had it not been for one extraordinary detail of the case: the crime happened not in real life but in a "virtual" hotel in the three-dimensional world Habbo Hotel, a children's game that only exists on the internet. (Keegan 2007: 16)

26. Although the analysis in this section will focus on *Second Life* (SL), there are other similar virtual worlds with similar options for the personalization of avatars, interactions and socialization. This is the case of *Blue Mars* (www.bluemarsonline.com/), *OpenLife* (http://openlifegrid.com/), *Inwordlz* (http://inworldz.com/), *Kaneva* (www.kaneva.com/), *ActiveWorlds* (www.activeworlds.com/), *Entropia Universe* (www.entropiauniverse.com/), *Twinity* (www.twinity.com), *Evolver* (http://evolver.com/), *Worlds.com* (http://worlds.com/), the Habbo hotel (www.habbo.es/), *Cybertown* (www.cybertown.com), *Dubit* (www.dubitchat.com/), *Moove* (www.moove.com/), *The Sims* (http://thesims.ea.com/), *Sora City* (www.soracity.com/) and *There* (www.there.com, now closed).

27. An extreme case of his reproduction of real-life attributes can be found in the film *Avatar* (James Cameron 2009), in which the protagonist (disabled in the physical world) literally *fuses* with the avatar that allows him to interact and move around the incredible Pandora jungle (another physical world in a distant planet). While the protagonist is locked inside a chamber, he lives an autonomous life inside the body of the avatar, and only when this intimate connection is interrupted by someone pressing a button outside the chamber (the avatar then collapses) does the protagonist return to his normal physical life as disabled. It comes as no surprise that, at a certain stage during the film, he comments that "everything is upside down," that his life in Pandora is what seems "real" to him, and that his normal life is strange to him.

> A British couple who got married after meeting in Second Life are divorcing after the wife caught her husband chatting up another woman in the virtual world. (Keegan 2008: 28)

> A loft in New York City. The singer Regina Spektor is performing songs from her new album. People wander in, sit down and discuss the music. Everything seems normal. […] The loft is a 3D computer animation […] it exists only on the internet. The audience is made up of virtual representations of real people. The real people sit at their computer screens around the world, living their lives through avatars, the characters that appear on the screen. Regina Spektor and her music are real people selling themselves in a virtual world. (D. Smith 2006: 13)

In general, SL offers an interesting intersection, imbrication and hybridization between physical and virtual life. We are offered a chance to escape from our boring ordinary lives and explore new identities. Nevertheless, as Boellstorff (2008: 120–121) stresses, we cannot easily escape from our physical lives. Actually, SL interactions frequently mould and define accurately the attributes that individual users already have in physical scenarios, rather than providing alternative lives or identities that do not overlap with their physical identities.

4.1 Terminological explanation

Virtual world is a label that I apply to a three-dimensional space such as SL, in which human figures or *avatars* exhibit a great capacity for nonverbal behaviour and interact with one another. As such, it is different from other similar Internet-mediated environments and thus a terminological explanation is required.

According to Bainbridge (2007), a virtual world is a simulated environment on the Internet that emulates the real world and whose inhabitants interact with avatars." Bartle (2010: 24) defines it as "an automated, shared, persistent environment with and through which people can interact in real time by means of a virtual self." Book (2004) underlines attributes such as a shared space, a graphic interface, immediacy, interactivity, persistence, socialization and a tendency to communal bonding. Besides, Hua & Haughton (2009: 889) stress the visual experience that virtual worlds offer users. SL possesses these features, and hence *virtual world* is an appropriate label for it. Another term, *virtual environment*, is similar but broader, since it also includes interactions between 2D avatars. And the term *collaborative virtual environment* is even broader, defined as "a digital system that allows geographically separated individuals to interact via networking technology" (Yee et al. 2009: 286), where avatar-mediated interactions are only part of the opportunities for interaction that the system provides.

SL and other virtual worlds are sometimes included in the broad category of *online games* and, specifically in the category of *massively multiplayer online role-playing games*, typically abbreviated as MMORPG, where the famous *World of Warcraft* is a paradigmatic example. If we compare it to SL, there are many analogies: in both there are 3D avatars interacting with one another and controlled by users, who display different nonverbal behaviours. However, in SL there is no pre-determined goal that, once achieved, signals the end of the game. Besides, SL recreates ordinary activities that cannot be labelled "playful" even if they are sometimes highly creative. In addition, many activities within SL resemble the ones performed in physical scenarios (Fetscherin & Lattermann 2007: 4).

It is also important to distinguish between *avatars*, created by users for their interactive goals, and *computer agents*, anthropomorphic visualizations of computer applications that are, therefore, of little interest for a pragmatic study of user-to-user Internet-mediated communication.[28] In any case, in the same environment there is very often a mixture of avatars (controlled by users) and agents (controlled by the computer system), as happens with many online games (Fox & Bailenson 2009: 148). In this case, these agents are usually labelled *embodied agents*.

Lastly, it is necessary to make a differentiation between types of avatars. Some studies do not differentiate between bi-dimensional avatars, with no mobility or capacity to generate nonverbal behaviours, and three-dimensional avatars that are fully animated. In this book I propose a distinction between *graphic avatars*, bi-dimensional and fixed, and *corporeal avatars*, three-dimensional and fully animated (the ones that we can find in SL). The definition of avatar by Bailenson & Blascovich (2004: 64) fits the kind of avatar that is interesting for pragmatics, the corporeal one: "a perceptible digital representation whose behaviors reflect those executed, typically in real time, by a specific human being." By contrast, definitions such as "general graphic representations that are personified by means of computer technology" or "graphic icons representing users through various forms," quoted in Vicdan & Ulusoy (2008), refer to graphic avatars. Although there is research on users' reactions and inferences when faced with these graphic avatars, it is more interesting for pragmatics (and *cyberpragmatics*) to study fully animated corporeal avatars such as the ones that interact in SL.

---

28. A possible interest for pragmatics would be, perhaps, to analyse the inferential steps that lead to an interpretation of the avatar created and controlled by a user and compare them to the inferential steps for the interpretation of the avatar-shaped computer agent. Nowak (2004), for instance, concluded in her study that there are no great differences between users' reactions to agents and their reactions to avatars, even in terms of credibility. Similar conclusions were drawn by von der Pütten et al. (2010).

At this point it also convenient to distinguish, as Boellstorff (2008: 133) does, between *primary avatar* and *alternative avatars*. The former is the one that the user chooses when he/she first joins SL. This avatar usually reproduces, with greater or lesser fidelity, the physical shape of the user. By contrast, subsequent *alternative avatars* are much more creative and even less anthropomorphic. These complement the *primary avatar*. Of course, this is not a norm that is invariably followed. For instance, a teacher at the University of Leeds comments:

> My original avatar is a furry and as different from me as possible. I deliberately went in [SL] as a fantastic creature unlike myself. I use this avatar for creative and social activities. My alt [alternative avatar] was created more recently so that I can participate in educational activities as myself. She looks a bit like me, though she is pale green, younger, and more glamorous. (pers. comm., November 24th, 2009)

4.2 Identity

The analysis of virtual worlds such as SL entails, again, the analysis of identity on the Net. The kind of identity display and shaping that we can find in this virtual world fits the term *liquid identity* proposed by Bauman (2005) or *hyper-identity*, as Adrian (2008: 368) calls it,[29] subject to specific goals that are changeable and made to fit different situations. These avatars may be mere additions to a solid physical identity, or authentic protagonists and the main sources of identity shaping for users who are not satisfied with their offline identities. These users find on the Net an escape from physical interactions and a transit to more *ad hoc* and satisfactory identities, created for specific goals (see Morie 2008).

Perhaps the most interesting research for an analysis of physical identities vs virtual identities in SL would be to analyse to what extent the physical ones influence the virtual ones and, especially, whether the latter influence the former. For example, a user (quoted in Adams 2006: 9) acknowledges how much his SL identity differs from his physical identity:

> [SL] hasn't changed much… but I certainly have. You need to be an extrovert to thrive in here. In RL [real life] I'm the opposite. I have a wife and family, but I spend a lot of time not really interacting with the outside world.

To explore other aspects of identity and overcome the limitations that one's body imposes on us in physical settings seem to be objectives that underlie SL

29. He uses a comparison between hyper-links and identities. In the same way as one is offered several links to click on when one accesses a web page, a user is offered several identities in virtual worlds to choose from and to fit specific communicative goals.

participation. Of course, avatar-mediated corporeal virtual identities do not have to resemble their physical counterparts. In this sense, McKinnon (1995, quoted in Kang & Yang 2006: 1175) suggests three forms of self-description of the user's identity in virtual contexts: *transparent expression*, when virtual identity and physical identity coincide; *translucent expression*, when virtual identity resembles more or less physical identity; and *opaque expression*, when both identities differ completely.

4.3 Body

An essential and obligatory step when joining SL and other similar virtual worlds is to select a body for the avatar, shape it in its main features and choose clothes for it. These choices inevitably generate inferences in other users, some of which will match the user's intentions, while others will be constructed beyond these intentions. This intended-exuded duality is of particular interest for a pragmatics of avatar-mediated communication. In theory, only the information intended to be communicated should be analysed by pragmatics, but valuable conclusions may be drawn from the avatar's exuded information. In this case, sex and body (and also clothes) stereotypes are usually at work in the form of mental schemas stored as part of "the culture" of a collectivity and which are inevitably reproduced in virtual worlds as intensely as in physical settings.

It is undeniable that when two individuals interact, the information about their sex, physical appearance and clothes help them frame the conversation correctly. In SL, the visual features of the interacting avatars also aid in choosing what kind of conversation is exchanged and which thematic course it takes (see Misoch 2008: 54–55). In fact, it has been demonstrated that the process of avatar creation in SL is intimately influenced by the user's (offline) bodily features, even if the user attempts to alter these features completely and create a radically different avatar (Vicdan & Ulusoy 2008), and even though avatar creation offers multiple possibilities for playing with identities and exploring other avatars' reactions.

Hence, even though users may try to create a radically different avatar from themselves, they cannot help bearing their physical bodies in mind when they choose features and qualities that improve the avatar in comparison to their physical bodies (e.g. broader shoulders, flatter belly, more handsome). As D. E. Jones (2005) qualifies, "virtual worlds feed societal fantasies developed within the mind/body discourse of transcending the deficiencies of human flesh. *Second Life,* which allows complete customization of avatar bodies, promises to give users a second skin that can improve on the corporeal and be changed like a suit of

clothes."[30] And the body maintains its influence in SL, as intensely as in the physical world. In Stromer-Galley & Martey (2009: 1051), for example, several studies are cited that reveal how users open themselves less to dialogues and keep greater distance with avatars that are not attractive, or how short avatars tend to exhibit less confidence in the dialogues with other avatars in which they participate.

The problem is that, either consciously or unconsciously, users often create the bodies of their avatars according to cultural or sexual stereotype-patterns of the community (in its narrow or wide dimensions). These stereotypes are stored in the form of mental schemas of great strength and easy access, and therefore require little mental effort in the process of the schema reproduction of "embodiment" of the avatar. Given the persistence and depth of the stereotypes of masculinity and femininity rooted in culture and enhanced by the mass media, it is foreseeable that these stereotypes, especially those related to the information *exuded* by the attributes of the body, will also be valid for the information provided by avatars in SL and other virtual worlds. And the same applies to sex roles. Guadagno et al. (2011) concluded that there is a parallelism between offline sex role expectations and the ones performed within SL. Specifically, they tested the *Social role theory*, which indicates that "men and women perform different roles in society with men primarily serving the role as provider and women primarily serving the role as caregiver" (ibid.: 305). Men and women learn different skills and beliefs that fit these roles and also impact their social behaviour. According to these different social roles, men and women are also subject to expectations for behaviour. These factors lead to gender differences in actual behaviour. The analysis of SL avatar behaviour yielded a confirmation of this theory, since

> overall, women reported engaging in more communal activities (e.g. meeting people, shopping) relative to men while using Second Life, and men reported engaging in more agentic activities (e.g., building things, owning and working property) relative to women. Furthermore, when describing their most positive experiences, women reported more communal experiences and men reported more agentic experiences. (ibid.: 307)

Besides, gender stereotypes are part of the encyclopaedic knowledge stored by the user and are updated in daily interactions which, based on relevance, cause certain assumptions to be strengthened or corroborated. The mass media are instrumental in spreading these assumptions, and play a substantial part bombarding the user with patterns of behaviour and models of corporeal identity. Some of these

---

30. Jones (ibid.) differentiates between *normative avatar bodies*, which behave in similar ways to human beings in the physical world, and *fantastic avatar bodies*, that hardly resemble human beings (e.g. furries).

assumptions can be transferred "vertically" through generations, while others are only valid "horizontally" within a community (although these limits are becoming increasingly blurred by globalization and the ubiquity of access to information from mass media and the Internet). In the second case (horizontal spread), and according to the *epidemiological model* envisaged by Sperber (1996), the transfer of information among humans is essential to make certain archetypical cultural assumptions more or less faithfully stored in the minds of all individuals:

> Most representations are found in only one individual, but some get communicated, transformed by the communicator into public representations and re-transformed by the audience into mental representations. Some even get communicated repeatedly, spread out in a human population and may end up being instantiated in every member of the population for several generations. […] Each member of the group has, in his or her head, millions of mental representations, some short-lived, others stored in long-term memory and constituting the individual's 'knowledge'. Of these mental representations, some – a very small proportion – get communicated repeatedly, and end up being distributed throughout the group, and thus have a mental version in most of its members. When we speak of *cultural representations*, we have in mind – or should have in mind – such widely distributed, lasting representations. (ibid.: 25, 33)

Gender stereotypes are examples of representations that spread across the population not only through interactions (e.g. parent-child) but also through archetypes that mass media reproduce and propose, and spread easily because the media have reached the status of "source of authority" for many people, especially teenagers. As a consequence, the audience tend to over-emphasize the importance of the body in both sexes, above other factors such as personality. It is not surprising, therefore, that these stereotypes are transferred equally to the three-dimensional world of avatars in SL.[31] This does not mean, of course, that all SL users share the same ideas on what model of masculinity and femininity (and the parallel communication and behaviour patterns by gender) are suitable for the physical and virtual interactions in which they participate, but even if the user does not share gender stereotypes, he/she is likely to be aware of the depth and extent of these

31. In Yus (2001c) a distinction was proposed between *woman-as-signifier* and *woman-as-signified* to refer to how women are portrayed by the mass media. In general, the former has been over-emphasized at the expense of the latter, hardly valued at all. This has generated a kind of *semiotic imbalance*. For example, the comedian Jo Brand (quoted in Wagg 1998: 122) stresses that "there is this attitude towards women which prevails in magazines and on the telly and if a Martian came down to earth and just had to watch telly and read magazines to find out what women were like he'd think that they were all blonde and 25 with big tits, you know. Because that is mainly what you expect on the telly. Also they would think that they were never rude and always looked nice, they always deferred to men, a lot of the time."

stereotypes. To account for this, in Yus (2002c) a distinction was proposed between *private beliefs*, those that the individual has acquired personally through interaction, communication and inference, and *metarepresented cultural beliefs*, those that the individual assigns as prototypical of a community or culture, that he/she assumes as widespread among its members, and that may or may not coincide with his/her their own private beliefs (or overlap in different degrees).

The bodies in SL, therefore, exhibit male and female archetypical qualities[32] and also prototypical clothing, all of them stored as mental schemas and transferred to the virtual environment from physical life (see Brookey & Cannon 2009). The same is true for other "inherited" traits of the body such as the race (see Harris et al. 2009). As Misoch (2008: 60) correctly points out, if the body is the instrument through which we exhibit our presence in physical spaces, the avatar takes on this role in the virtual world, while performing the function of being a visually perceptible and stable receptacle of identity. Furthermore, these stereotypes are often increased in SL (e.g. disproportionate chests in male avatars and big breasts[33] in female avatars), taking the stereotype to extreme lengths. It is not surprising, then, that De Salvador Agra (2009) concluded in her research that in SL

> there are no socially undervalued virtual bodies. In this supposedly simulated world, where almost anything is possible, we can not find any avatar to take the special feature of being disabled, fat, lame, squint or any other typical feature that is object of discrimination in the old offline world. That is, in our incursion into this info-virtual environment, we could see how nobody in this freedom that supposedly defines the environment, opts for a body that is pejoratively valued in our physical society.

---

32. In general, the sexual stereotype of women in SL is "thin, narrow hips, long legs and generous breasts." The male sexual stereotype is "an athletic build, broad shoulders, muscular chest and narrow hips" (Misoch 2008: 59). These correspond to the images that are incessantly communicated to the audience of Western cultures from the mass media.

33. In a survey by the blog *Pixels and Policy* (November 2nd, 2009), 70% of female users surveyed admitted that the design of their avatars' breasts was one of their main concerns during the design of their avatars. Some users showed a clear submission to the woman's sexual stereotype of "exuberant" as a means to be desired and obtain more frequent interactions with other avatars, as can be concluded from the testimony of one of the female users: "At first I played with an avatar that I thought represented me physically, but not many people talked to me. Now [with a big-chested avatar] people go out of their way to IM me and send me friend requests." According to that blog, this is a clear setback for female users, who place the passive and stereotypical role of attractiveness before actively seeking friendship based on personality. Note, however, the study by Ducheneaut et al. (2009), in which informants stressed the importance of the process of hair design for the avatar, over the process of design of sexual attributes.

One consequence of this transfer of body stereotypes from the physical context to the virtual context of SL is what can be called *illusion of corporeal avatar*, that is, the illusion that it is the body of the avatar, and not the user who created it, which holds all the communicative activity in SL interactions. Bente et al. (2008: 134) also comment that "although aware of the avatars' artificial nature, users seem to respond to their appearance much in the same way as they do to humans in real-life encounters." This illusion would explain why many avatars are treated according to their virtual appearance, even though it is obvious that the actual appearance (or even the sex) of the user can vary dramatically compared to that artificial appearance.

De Salvador Agra (ibid.) entered SL with three avatars of radically different appearance: a young and voluptuous woman, a bald old man and an obese black woman (Figure 5.1). Although the user is the same and her physical appearance has nothing to do with these three avatars, she suffered indiscriminate abuse when she used the avatar of an obese black woman, jokes and derision when she played the part of a bald old man (specifically its avatar was heckled by a group of avatar-girls that called him "grandpa" and asked him what an old man like him was doing there) and, finally, sexual approaches when incarnated in the body of a young attractive woman. In other words, participants completely ignore the fact that there may be little or no resemblance between the avatar and the user who created it and, instead, the body shape of the avatar is idealized as an inherent aspect of SL interactions. In similar terms, a professor at Ball State University (Indiana, USA) comments that

> Of course on an intellectual level I realize that the avatars are fictitious, probably quite different from the people who created them, and that there is some real person sitting at some keyboard somewhere in RL [real life] but all of that recedes way into the background while in SL. I identify to a surprising degree with my avatar; it feels as if I'm her, actually visiting places, doing things, and talking to people (that is, talking to avatars, but I don't really relate to them as avatars while in SL; I seem to view them as people). When someone I've been talking to turns out to be a vampire, I feel a little worried, as if the situation were much more real than just images on a screen. (pers. comm., November 26th, 2009)

According to Yee, Ellis & Ducheneaut (2009), this phenomenon is part of a more general tendency: the human desire to reproduce the parameters of the physical world in the virtual setting and, in fact, for these authors SL and other virtual worlds do fit, to a greater or lesser extent, several expectations of duplicity: (1) the *expectation of human embodiment* (to use human-resembling avatars); (2) the *expectation of matched affordances* (the avatars move and do the things that humans typically do in physical contexts, but in SL the avatars can fly and tele-transport

themselves to other spaces of this virtual world); (3) the *expectation of congruence* (users, by means of avatars, possess different conceptualizations of what a virtual world is, but these are congruent); and (4) the *expectation of a single avatar control* (each user can control only one avatar at a given moment and each avatar is controlled by a single user).

Figure 5.1 Avatars of De Salvador Agra (2009)

4.4 Verbal interaction

There are three main ways to interact verbally in SL: the chat room option, the instant messaging option (IM) and the paid-for voice service.[34] The first two are fairly common in other virtual worlds, but there are also significant differences with pragmatic implications. Thus, when users type their messages in the chat or IM window, the server reproduces the messages on the screen in their entirety after being sent by the "addresser user" (i.e. not word by word as the user is typing it), in the same way as in conventional text-based chat rooms. One cannot, therefore, be generating inferences and interpretive hypotheses as the text is appearing on the screen. These inferences may be corroborated or refuted with the processing of the next part of the message that appears on the screen (thus forcing the reader to backtrack and re-interpret the text according to the "new evidence"

34. Pujolà & Palomeque (2010: 136) summarize most of the forms of interaction in SL: (a) *local chat*: text-based, synchronous and public; (b) *voice chat*: voice-based, synchronous and public; (c) *IM*: text based, synchronous (also asynchronous) and private; (d) *voice call*: voice-based, synchronous and private; (e) *notecard*: text-based, asynchronous and private; and (f) *gestures*: non-verbal communication (the user has a default set of gestures in the inventory but can also create new ones for their avatar).

provided by the text just typed). In SL, while the text appears on the screen, the avatars automatically mimic the action of typing text with their hands on an invisible keyboard. Moreover, if the avatars come within 30 (virtual) metres away from other avatars, then they can "hear" what they are saying, that is, the user can read the conversations that other avatars are participating in if the user draws the avatar close enough.

In the virtual world *There* (recently closed), the protocol of verbal interaction was somewhat different, since it was based on a number of balloons, as in comics, displayed above the avatar's head (when there were multiple utterances, these were inter-connected with each other forming a chain of balloons) and in which the text was being copied as the user was typing it. This quality made it possible for interactions to have overlappings or interruptions that are nonexistent in SL. In addition, the users were able to change the content of their messages while they were reading the other avatars' messages as they appeared in the balloons, and hence the construction of the message obtained a high level of relevance and appropriateness in the context of that conversational interaction.

Of course, in the text typed both in the chat and IM windows, users reproduce the strategies of compensation for the absence of orality and of text deformation that have been discussed earlier in this chapter (e.g. repetition of letters or punctuation marks, capitalization, use of emoticons). However, the use of one or another channel has different connotations, as noted by Boellstorff (2008: 152–154). In general, SL conversations between avatars that have just met are usually held in public using the chat room facility. But it is possible to add an avatar as a "friend"[35] and at this moment the private IM is preferred, especially with friends or contacts who know each other well, just like in conventional IM. Very often, and even if there is no other avatar around, users leave the chat room and continue the conversation through IM when the topic of the conversation becomes more confidential, private or intimate.

Similarly, the use of multiple windows for interactions in parallel is common. The control of these windows, together with the management of the nonverbal behaviour of the avatar, can generate extra effort in SL communication. This effort can be increased by the possibility of engaging in chat room and IM interactions simultaneously. Another extra effort is required to manage the accumulated IM

---

35. The procedure for doing this is to click on the avatar that one wants to add as a friend with the right button of the mouse and choose "add friend" from the context menu displayed after the click. On the user's screen that controls the behaviour of the avatar the message "X is offering friendship" then appears where X is the name of the avatar, which is always visible above the avatar in SL. If the user enters "yes," from this moment the added user will always be able to locate the avatar in SL and will know when it is online.

messages when the user logs onto SL. In fact, one reason why many users create *alternative avatars* is that it enables them to log on anonymously and not be bombarded by IM messages.

In the case of using the voice option, users definitely get an extra level of contextualization that allows users to "frame" the interaction and identify the underlying communicative intentions efficiently. This is a paid-for option and is not selected by most users, many of whom are satisfied with the options provided by text-based interaction (see Bente et al. 2008: 288–292), as in the case of the researcher at the University of Leeds mentioned above:

> [My avatar] has used voice. I prefer text, especially in my original avatar, as it allows me to be someone else. I also like the ability of text to neutralise social features, such as gender, region of origin, social class and so on. I think it empowers people more than voice and is easier across national boundaries. So I don't think it is a limitation. My avatars can express emotion through actions and expressions and through statements of actions and feelings.
>
> (pers. comm., November 24th, 2009)

In Boellstorff (ibid.: 114), the opinions of several users on the voice option are quoted. For many of them, voice provides a high level of intimacy, while for others the voice "kills the fantasy" of true avatar-mediated communication in SL. For one user, refusing to use the voice was interpreted as proof that she was actually a man. For another user, the voice in SL is a natural extension in SL of interactions that take place in physical contexts. But this same argument is considered negative for others, because it ruins the illusion of interaction among avatars.

4.5 Nonverbal behaviour

In previous pages I mentioned the example of a person yawning and the possible interpretive outcomes depending on the axes of "intentional / unintentional" and "understood correctly / incorrectly." SL avatars are endowed with a great capacity to generate nonverbal behaviours. In theory, it should be clear that these necessarily have to start from an intention by the user that generates them in the avatar. However, as discussed below, the qualities of the computer application that manages SL can lead to interesting situations in terms of avatars' nonverbal behaviour.

Nonverbal behaviours, beyond the basic movement or simple behaviours of the avatar, are usually generated by applications that are available in the environment where avatars interact. For example, in a dance floor context two users can click on the button and select from the menu an animation that gives avatars the ability to dance. If two avatars are getting married (which occurs quite often in SL) and an avatar-priest says "you can kiss the bride" each user clicks on the

animations menu and chooses "kiss" so that their avatars can carry out this nonverbal action. In *There*, these animations appeared as blue circles around avatars, also called *action tags*. One clicked on the tag and a menu of possible actions turned up (e.g. "sit next to the nearest avatar"). In *There*, it was also possible, for example, to type a prototypical nonverbal behaviour (e.g. smiling, yawning, blushing) and automatically the corresponding nonverbal behaviour was generated in the avatar. In fact, the nonverbal behaviour available in that virtual world was quite advanced, since the avatar could, for example, sip a drink or offer it to other avatars.

Managing the full range of possibilities regarding avatars' nonverbal behaviour involves effort and training for which not all users find a reward in communicative terms (cognitive effects). For example, Suler (2007) describes his tiring training in the management of the nonverbal behaviour of his avatar:

> It took me several minutes just to figure out how to move my avatar, and then I was literally walking into walls and trees. I spent most of the first day learning how to move about [in SL] without looking like a complete idiot, how to visually survey and interact with the environment, and, most fun of all, how to fly like superman. The controls for navigating one's avatar are much more sophisticated than they used to be at the Palace. This posed a rather interesting challenge. Even after several hours, when I thought I was doing reasonably well, a more experienced user who I met in the SL version of Amsterdam commented on me being a newbie. When I asked how she knew, she replied, "By how you walk."

As noted above, managing the nonverbal behaviour of the avatar entails an underlying intentionality by the user that the avatar should behave in a certain way so that, in principle, we could not find in SL all the unintentionally *exuded* nonverbal behaviour that so often helps people understand each other correctly in physical settings. For example, a prototypical exuded behaviour like *blushing*, uncontrollable in physical contexts, must be generated by the user, and thereby loses much of the "communicative naturalness" that SL aims at, and forces the users to constantly assess the type and intensity of their feelings and emotions. To this we must add the difficulty of making a computer program perform the full range and variability of nonverbal behaviours, which is evident, for example, in the case of facial gestures, as will be briefly discussed below.

However, there are situations in which "unintended nonverbal behaviours" are produced in the avatar. For example, the software can make the avatar behave in a strange way, sometimes due to the user's inexperience and occasionally due to alterations in the program's response to the interactive goals of the SL residents. These situations introduce the possibility of an unintended or *exuded* nonverbal behaviour from avatars. Therefore, for any type of nonverbal behaviour in avatars

it is important to note the role that the interface can play in the outcome of these behaviours along the intentional / unintentional axis. For example, Antonijevic (2008) proposes the following classification that in my opinion satisfies this requirement:

1. *User-defined cues*. These refer to behaviours that the user deliberately performs, such as separating the avatar from another avatar that is too close by using the keyboard arrows. These cues clearly play a part in interactions, because they reveal communicative intentions and regulate interactions.

The most common behaviour in this first type is the management of proxemics between avatars, that is, the control of the distance between avatars and the interpretations that this distance conveys. In fact, as concluded in Bailenson & Blascovich (2008: 2676), the avatars in virtual worlds like SL often behave in the same way as human beings in terms of interpersonal distance. While there are no "written rules" about what personal distance is the most appropriate in each case, users-avatars consistently identify the most appropriate distance and explicitly show their disagreement when it is not respected by another avatar (Taylor 2002: 42) .

2. *Predefined cues*. They refer to nonverbal behaviours that the user has not generated in the avatar, but it is the computer system that produces them. For example, if the user stops typing text in the chat room or the IM window, the SL system detects user inactivity and forces the avatar to adopt a particular position. Something similar happens to what Boellstorff (2008: 106–107) calls *away-from-keyboard problem*. Users frequently stop controlling the behaviour of their avatars and leave the computer for a while without actually turning off the system, or they stay at the computer but working on another application in parallel. In both cases, the system causes the avatar to adopt a position of complete inactivity and to appear to be dead on the screen, to the astonishment of the other avatars. In theory, after several minutes, the system detects the "sustained inactivity" of the avatar and makes it disappear from the screen, but there are programs available on the Net that invalidate this action.

A sub-group of these predefined cues includes behaviours that seek to regulate and synchronize the interaction among avatars. For example, if a user moves his/her avatar in a certain direction, the heads of nearby avatars automatically turn in the same direction, in an attempt to reproduce the conditions of interaction that take place in physical contexts.

3. *Blended cues.* This category refers to nonverbal behaviours chosen by the user, but carried out by the computer system of the virtual world. All predefined animations displayed in the contextual menu after clicking on the right button of the mouse are included in this category. For example, a user may want his/her avatar to kiss another avatar, but the kiss itself is made possible by an animation

that the system displays, without intervention from the user. In addition, this type of behaviour is heavily influenced by the context and the sex of the avatars. For instance, the action of sitting down is different in animations for male and for female avatars and it is likely that, more or less consciously, archetypical patterns of behaviour according to the person's sex may have been followed (in the form of accessible stereotypical mental schemas) in the design of many of these animations.

4. *Missing cues*. Finally, Antonijevic (ibid.) mentions human nonverbal behaviours that the system is unable to reproduce, although constant evolution in the various versions of the software in these virtual worlds leads to the incorporation of new behaviours in the range of nonverbal behaviours of the avatar. For example, the addition of a voice in SL has been a major change and users are able to communicate, with their avatars' words, the full range of vocal nonverbal behaviours available to humans in face-to-face interactions.

On the other hand, of all the areas of nonverbal expressiveness, it is the human facial expression that probably best illustrates the problems of these virtual worlds to mimic the nonverbal behaviour produced by humans in physical contexts. Currently, at least with the technology available, it is very difficult to reproduce accurately the full range of emotions and feelings that are based on facial gestures, and we can often see that only very general behaviours such as smiling appear in the catalogue of avatar animations. For instance, Donath (2001) stresses that it is not simply a matter of getting more detail in the avatar's facial expression, but there are also many cognitive and cultural determinants that influence the use that humans make of the expressions on their faces, and these are very difficult to transfer to the virtual world. For her, introducing the face in these environments involves a radical reinterpretation of what we consider "personal appearance" and, at the same time, trying to maintain the cognitive and cultural meanings that we associate with the familiarity of the facial gesture. Donath (ibid.) does not consider that avatar expressions should be a duplicate of those found in physical spaces and, in fact, the virtual environment offers many possibilities beyond these natural gestures. However, she considers that we have to be careful when facial gestures are introduced in these environments, because they are full of subtle information, and a poorly designed expression can lead to misunderstandings and cause virtual interactions to fail.

This issue is accentuated by the role of the face in the regulation of conversations. Indeed, facial gestures are *regulators,* in the sense proposed by Ekman & Friesen (1969), of conversations with speakers providing evidence that they are paying attention or sharing an opinion. This role has also been transferred to virtual worlds. In the study by Garau et al. (2001), they concluded that having an avatar whose facial behaviour can be directly related to the development of the conversation increases the quality of virtual communication, compared to those

conversations where avatars exhibit random gestures. Put simply, it is not enough for the avatar to be "extrovert" with other avatars, it has to engage facially in the conversation in order to be successful.

However, advances in computer applications that manage interaction in SL and other virtual worlds open up interesting possibilities with clear pragmatic implications. For example, Heike et al. (2009: 207) mention studies that attempt to get the avatar to nod upon detecting the end of the interlocutor avatar's utterance (i.e. to identify the *turn transition place*, in the terminology of conversation analysis). Even more interesting is the advancement cited in Morrás (2009) about the invention of an SL viewer that permits the transference of the user's emotions and gestures to his/her avatar. Therefore, the avatar makes in real time the same gestures and movements that the user performs. This is a very interesting evolution and opens up new areas for pragmatic research in the future, because this interface provides the avatar with greater capacity for facial expressiveness. These advances, together with advances in general avatar design options influence communicative outcomes in SL. For example, Hussain et al. (2011) concluded in their study that there is a high correlation between humanness of avatars and high credibility measures. Therefore, a good design of the avatar may have an important impact on interpretive outcomes and the overall SL user's satisfaction.

Finally, it should be noted that many nonverbal behaviours do not have an inter-cultural applicability. Users come together in a three-dimensional environment that is accessible from anywhere in the world that has Internet access. It is predictable, therefore, that there may be misunderstandings due to ignorance of the nonverbal behaviours that are typical (and taken for granted) in the culture to which the user who is handling the avatar belongs. This is very likely to occur if we consider that on many occasions there are no cultural clues in the final design of an avatar and the real user behind the avatar may be managing its nonverbal behaviour from anywhere (and the corresponding culture) in the world. For example, Koda et al. (2009) compared the expressions of the avatars designed by Western and Japanese users and which depict typical faces of these cultures. The investigation concluded that, in fact, there are cultural differences in interpreting the facial expressions of avatars. It was also concluded that facial expressions of a positive nature have a greater cultural variability in their interpretation than those of a negative kind. Specifically, the subjects' interpretations of negative expressions (sad, disapproving, angry, and confused) were similar to the designers' intentions, regardless of the country, that is, the subjects' answers to those expressions were similar across countries. On the contrary, the subjects' interpretation of positive expressions (happy, approving, proud, grateful, and impressed) varied across countries.

5. Videoconferencing and context accessibility

If we arrange all *cyber-media* on a scale of options for contextualization and capacity to convey vocal and visual nonverbal information typically communicated in face-to-face interactions, videoconferencing would no doubt be ranked highest in contextualization, since it closely resembles face-to-face dialogues even if framed by the computer screen.

It is convenient to start with some terminological clarification. Mouzourakis (1996: 22) proposes a terminology to distinguish between (a) *teleconferencing* (any form of communication that involves the use of at least one audio channel, between two participants in a meeting but separated by some distance), (b) *audioconferencing* (teleconference in which the sound is the only channel of communication), and (c) *videoconferencing* (a special kind of conference in which the image is used in addition to the sound). We should also distinguish between *video-mediated communication* and *videoconferencing*. The former involves the use of television cameras that transmit the signal to a satellite which, in turn, distributes the signal to other parts of the world. In *videoconferencing*, by contrast, the signal is transferred through the Internet in digital format and therefore can be analysed from a *cyberpragmatic* point of view. However, the threshold that distinguishes the two forms of signal transmission is becoming diffuse, as broadcasters often use the Internet to distribute video content and rely increasingly on satellites and wireless networks for transmission.

It is noteworthy that achieving close fidelity in terms of contextualization does not necessarily mean that users' estimation of relevance will not be affected by the screen-framed transmission of information. Indeed, the first problem in trying to equate videoconferencing and face-to-face interactions lies in the fact that videoconferencing links spaces that can be very distant from each other, but these spaces are framed by the computer screen. Sometimes, alterations of relevance can be produced due to the mutual influence of the framed area of the room and the one which is out of frame. This is especially clear in the case of dialogues among people who share the same physical space of a room but only one of them is framed inside the screen and, at the same time, all of them in the room have a dialogue with another user through videoconferencing (Raudaskoski 2000). Fayard (2006: 154) also comments that the introduction of video in the interactions produces asymmetries that affect everyday communicative practice, thus making the routines and scripts of interactive behaviour (which we use almost unconsciously) less valid. In videoconferencing, the interlocutors do not share the same physical context, which affects the interpretation of the environment of the interlocutor, his/her eyes, body expression, etc. The social context is not shared, either, and that generates a loss in the appropriate contextual clues to interpret the

speaker's communicative behaviour. Besides, some technological problems have to be taken into consideration which disrupt communication (Have I heard correctly? Has the interlocutor grasped the meaning of my gesture? Is there synchronization between voice and image?).

Furthermore, the concept of *stage*, in the sense proposed by Goffman (1987), should be emphasized. I referred to this term at the beginning of Chapter 2, when I stated that we must distinguish between the roles we play in society at the *front stage* of interactions and the personal reality that lies at the *back stage* of our identities, the part that hides behind this "social playground." These roles also play a part in videoconferencing. What proportion of the roles that are filmed by the camera belongs to the "social front stage" and how much refers to the real identity of the interlocutor?

In short, videoconferencing offers a fully contextualized environment that conveys users' vocal and visual nonverbal behaviour but suffers from some problems caused by the lack of physical co-presence, screen-framed interaction, and mediation. These problems are more or less obvious depending on how faithfully the filmed scenario is reproduced in the transmission of different utterances exchanged through this type of Internet-mediated communication.