MARIA GRAZIA SINDONI

DIGITAL APOCALYPSES: VERBAL OR VISUAL DECLINE?

This paper addresses the question as to whether new digital forms of interaction instantiate a decline of verbal language or of visual communication. Many studies claim that web-based interactions are more based on verbal language, as they involve live, spoken interactions, which are by definition shaped by conversation, whereas others argue that image is dominant in the digital domain, as it deploys affordances typical of multimodal environments.
It provides a theoretical introduction to the study of videochats, seen as dynamic texts that enact a new form of interaction. Interactional resources are arranged in such innovative ways in videochats that they produce a new field investigation for researchers in the field of linguistics, communication, media studies, visual ethnography and digital literacy. The paper singles out relevant aspects, such as the alternation of speech and writing, new patterns in proxemics and kinesics, impossibility of eye contact, which are so striking as to warrant special investigation. In particular, speech and writing are technologically integrated, allowing participants to modeswitch, i.e. alternate between spoken and written discourse. New arrangements of verbal and non verbal resources attempt to simulate face-to-face conversations. However, the illusion of a face-to-face conversation dissolves as soon as videochat-specific resources are unpacked. Despite the growing research into non verbal behavior, videochat data challenge visual analysts and researchers for a variety of reasons illustrated in this paper.

Digital apocalypses

Celebrators of the Internet heralded it as a technology which placed cultural acts in the hands of everyday users, decentralizing speech, publishing, film-making, radio and television broadcasting (Poster 1995). Transcending the limits of the passive reception typical of the broadcasting era with Internet interaction was announced by second media age literature as a major breakthrough (see Reinghold 1994 for communitarian approaches and Guattari 1989 for postmodern theories). Second media age literature also contends that face-to-face interaction is superseded by extended forms of communication, such as the Internet. In this view, technology separates the individual from a "natural state" of interaction, i.e. face-to-face interaction that encourages other forms of social interaction and integration. The latter is endorsed through the erasure of ethnic and gender differences (Poster 1997). However, others argue that since face-to-face interaction is so important as a means of connecting people up in the information society, Internet has, as a consequence, become all powerful and all pervasive as a means of instantiation of such connections (Holmes 2005). Holmes, in fact, claims that communication environments *frame* individual lives, regardless of individual communicative acts: "the dominant background connections or mediums by which a given group of individuals are socially integrated come to mediate other levels of interaction" (Holmes 2005: 17). Holmes argues that our times are characterized by a mode of communication which is

no longer connected to a specific medium. The consequences are twofold: a feeling of liberation and greater freedom to communicate − but also the establishment of new social bonds. In other words, whilst experiencing a telephone call, we may perceive it *as if* we were involved in face-to-face conversation *even though we are not*. But we may also have a *half-present* face-to-face conversation, as extended forms of communication affect *what* we experience face-to-face and *how* we do so (Holmes 2005).

Predominant forms of social integration mediate communication in our age. Neil Postman's media ecology has emerged from McLuhan's influential medium theory which focused on the medium, not as a material channel conveying messages, but rather as determining the nature of the message itself. Whereas McLuhan famously claimed that "the medium is the message" (McLuhan 1964), subsequently arguing that "the medium is the massage" (McLuhan and Fiore 1967) and further claiming that anything that can extend the body's senses and capabilities may be defined as a medium, Postman (1985), on the other hand, views media in terms of environments, shaping and being shaped by the context. In a similar vein, Wesch (2009) claims that communication media influence the nature of our interactions, ultimately determining our attitudes, opinions and stance.

However, this discussion is not limited to human-related interactions. We need to broaden our horizon of investigation and move up another notch. Another level of analysis helps understand more easily that what we are witnessing is a profound bio-anthropological change. As biological organisms, we human beings are mutants. We are currently in the process of altering our strategies of interaction as a result of new tools. In other words, the incidence of mediation (i.e. presence of technology) in "mediated communication" has reached extraordinary levels of stratification and complexity that it is paving the way for new forms and norms of interaction (Scollon 1998, 2001). The new emergent texts that modern technology brings about allow oral and written modes of communication to be intertwined. This has sometimes also been the case in the past, most notably with the invention of printing.

The history of human communication is punctuated by micro and macro biological and socio-historical events that amount to changes or fractures, altering the linear course of human history and signaling predicaments or conjunctures that require the use of new epistemological tools to study and make sense of them. The development of visual/graphic semiotic systems pinpoints in time how different cultural stages in human history have been articulated and negotiated. The interpretation of such systems is neither optional nor limited to the development of writing systems.

Does our society simply require modes of communication that are more complex and sophisticated than those of the past (Kress 2003, 2010)? Or is it the case that the human body is changing and that the robotic extensions perceivable (e.g. mobile phones, netbooks) are part of the evolution of man's body?

In the 1980s, Ong (1977, 1982) discussed the technologies of the word in terms of the distinction between orality and literacy, arguing that the invention of writing caused a deep and unavoidable process of changes in the human brain, ultimately modifying our capacity for storing information, and transmitting and distributing it. What was previously focused on the immediate context of communication, relating to immediate needs, was later turned by the technology of writing into the possibility of analytical evaluations of the context itself, whence the rise of complexity. In other words, oral cultures do not have documents, but memories. People who are bred in oral cultures learn, but *do not study* (Ong 1982).

What has all this to do with systems of communication developed within the domain of what may be defined as part of the evolution of the web? The answer is that Web 2.0 is a stage in the evolution of the web, encouraging peer interactions, lower thresholds in the distribution of information, integration of technologies and the development of new systems of human socialization (i.e. social networks, videochats, video sharing communities, etc.). In other words, Web 2.0 allows new ways of interacting with people, sharing, selecting and discarding information according to one's own taste or opinion, in a post- or hyper-modern version of pop culture, where "copy and paste" techniques are the implied and never-questioned rule and where intertextuality is so dense and saturated as to become nullified.

New technologies pave the way for new forms of interaction and with it a world of paradoxes, illusions and deceptions. Below a specific text genre is explored in this deceptive, illusive and paradoxical light: even as it mimics face-to-face conversations, the videochat distances itself from *in presentia* interactions.

Speech and writing in videochats

What is a videochat? It is a form of computer-mediated communication (Herring 1996) connecting two or more people from different places through online, simultaneous two-way audio and video transmission. It is widely used for work, personal, and business communication. But how is it relevant to the study of speech and writing in web-based communities?

In videochats, participants may use speech and writing at the same time, conversing but also writing and sending text comments. They *may* and in fact often

do modeswitch, i.e. alternate between speech and writing and, in so doing, they are able to mix both modes in personal ways and achieve specific communicative purposes by drawing on both resources at the same time (Sindoni 2010). Modeswitching is a descriptive and exploratory label which takes this communicative possibility into account. It paraphrases the linguistic notion of code-switching, which is generally held to be the use of more than one language or variety in a conversation (Heller 1988). It also draws on Halliday's concept of written and oral modes of communication (Halliday 1978; Halliday & Mathiessen 2004) and is based on the observation that Web 2.0 has so far developed a wealth of new texts which intertwine speech and writing in unparalleled ways. Videochats basically rely on speech technologies, but there are constant alternations between text chat and audio/video conversations, especially in more informal contexts. In other words, an online conversation may be interspersed with written comments.

Preliminary research questions include: what drives users to associate oral and written technologies of the word – to borrow well-known categories from Ong (1977, 1982) – in videochats? Is it because they are just there or is there something else lying behind, or perhaps beyond, the mere exploitation of the full technological affordances available in videochats? More than two decades ago, Halliday argued that the categories of "written" and "spoken" are indeterminate. Hence the need to specify which discourse variable is being referred to when a given mode is selected. As a working hypothesis, in the case of videochats, the videotrack may be considered as the main online conversation, while written comments might be defined as integrations of the main text. However, many instances of videochats can be defined as the exact *opposite*: written comments constitute the main text, interspersed with occasional video/audio comments. The great variability in these integrations needs to be further explored within a purposely-created but empirically validated theoretical model.

Halliday identified distinct arrays of features for oral and written modes, highlighting the fact that both modes are equally systematic, organized, regular and productive of coherent discourse (Halliday 2002 [1987]: 340). Given the prominence of grammar for writing (theories on grammar were written in books, and, as such, typically based their models on written discourse), Halliday was at pains to create theoretical bases for a grammar that specifically took spoken resources in the linguistic system into account.

While speech exhibits higher levels of grammatical intricacy, writing displays a higher lexical density. So Halliday claims that both modes are equally complex. Instead of trying to indicate the primacy of one mode over the other, the question is how to define and identify kinds of complexity, occurrences and degrees of

overlapping between the two modes (Tannen 1982, 2007). Speech and writing are blurred notions and this is particularly true in the context of spoken and written informal contexts vs. spoken and written formal contexts. To bypass this problem, Lakoff proposed the distinction between spontaneous vs. non-spontaneous discourse in the early 80s (Lakoff 1982). However, digital systems of interactions, as videochats, further blur this useful distinction when both modes are used at the same time and in the *same* interactional event. Thus, the notion of modeswitching becomes a central theoretical concept to tackle videochats.

In the examination of clause complexes, for example, concepts such as embedding and hypotaxis need to be analyzed separately. Halliday considers this strategy as fundamental "to do justice to the particular mode of organization of both spoken and written discourse" (Halliday 2002 [1987]: 343). He also points out that: "grammar needs to distinguish between the constituency relation of embedding, or rankshift, where one element is a structural part of another and the dependency relation of 'taxis', where one element is bound or linked to another but is not part of it" (Halliday 2002 [1987]: 343-4).

In other words, only a grammar which considers cultural, social and linguistic distinctions between speech and writing fully accounts for differences and overlaps pertaining to both modes (Halliday 2002 [1987]). A comparative analysis of the two different types of texts, oral and written (audio/video text and written comments), serves the purpose of monitoring and including *all* possible convergences and discrepancies as well as decoupling, given that writing during a videochat will show features characteristic of speech, such as "spun out", "flowing", "choreographic texture", etc.

Videochats simulate face-to-face conversations, giving the illusory perception of sharing the same context of situation, a visual and psychological perception which is created virtually through a range of different strategies. A fundamental factor adding to the illusory perception of a face-to-face conversation is that verbal exchanges occur in real time. Moreover, users see each other in real time, so that conversations reproduce long or short verbal exchanges *in presentia*, with turn-taking, pauses, hesitations, etc., respecting social roles and cultural conventions. However, this system of interaction is a rather crude imitation of what happens in face-to-face conversations. First and foremost, participants do not share the same context of situation. Just as in face-to-face conversations, they take turns; however, they are physically separate and what they see is a projected image of their discourse partner. Videochat software systems commonly provide a larger-scale close-up of the discourse partner located in the foreground, while one's own smaller-scale close-up image appears at the bottom or at the top of the frame.

Moreover, the size of one's own image and that of one's discourse partner varies from one software system to another. Interest is mainly focused on the discourse partner, but one's own image is nonetheless there, making oneself aware of one's own body movements, posture, etc.

From an intersemiotic perspective, videochats include two differently integrated frames showing live images of users. Contrary to what happens in face-to-face interactions, each discourse partner has the unique opportunity to see himself/herself while having a video conversation. Observing oneself during a conversation produces a series of psychological effects influencing the verbal and non verbal characteristics of the online exchange. In other words, casting a sidelong glance at oneself during a conversation may change, if not determine, the way one speaks, gesticulates, smiles and so on; it entails adjusting to the addressee's expectations, but mostly adjusting to one's own ever-changing expectations as well. Expectations also have to do with *how* we interact, using our interactional resources to create the conditions for successful communication. In the next sections, relevant resources are discussed to expand our notion of videochats as a digital genre.

Representing social distance

Perception, use, structuring and management of space all have the effect of framing social expectations, interactions and ways in which actual communicative events take place. Social interactions, at least when defined by face-to-face communication, occur in a setting, which is physically perceived by participants, who negotiate their relationship with space and with one another in meaningful patterns. Hall (1966) defined this area of interest "proxemics", distinguishing four different spatial distances (i.e. intimate, personal, social/consultative, and public).

Social distance in videochats is fixed, and participants stand in a frozen space. This means that medium constraints (close-to-very-close shot of the projected image of both participants) do not allow mutual negotiation of space. To be more precise, distance is not established by those who interact, but between one participant and one machine. This distance foregrounds the *representation* of distance between users. Since space reflects environmental arrangements that may encourage communication (i.e. sociopetal space, Hall 1966) or keep people apart (sociofugal space, Hall 1966), we may wonder whether a videochat fits into the first or second category. Furthermore, studies on proxemics have come to include postural identification (i.e. sitting, standing), distance, frontal orientation, and body positioning (Harrigan 2005:

149) which in this case are not always easy to discern, especially if they are partially visible or fall outside the projected frame.

Ways of managing space during multimodal interactions affect the ways in which communication is organized overall and carried forward (Heath & Luff 2000). Notions of territoriality may well refer to the act of laying claim to, and defending, a territory from political, ideological, cultural invasions, but may also refer to personal spaces, whose "invasion" may be perceived as equally, if not more, threatening. Hall and Hall (1990) argue that space is perceived by all senses, not by vision alone.

However, proxemics may also be interpreted in broader terms, for example taking into account how historically people have come to negotiate space with other people in public places and how notions of personal space have shaped, related or opposed notions such as those of neighborhood, solidarity, estrangement and alienation. Simmel was one of the first to comment on how social intercourse in modern times is more based on the sense of *seeing* than on the sense of *hearing* (1921 [1908]). Furthermore, sociability takes on different articulations in urban and rural areas. Cities are considered hostile and unwelcoming places where anonymity is the rule. According to Lofland (1973), coping with the city means coping with strangers.

Goffman defined the personal space as "space surrounding an individual, anywhere within which an entering other causes the individual to feel encroached upon, leading him to show displeasure and sometimes to withdraw" (1971: 29-30). Social interaction is thus intertwined with proxemics broadly interpreted as a meaningful way to establish and maintain relationships. The body helps achieve this communicative task, for example by directly modifying it (e.g. tattoos), by adorning it (e.g. clothes), by moving it (e.g. movements, posture, etc., Emmison & Smith 2000: 212.)

In videochats, one's personal space, and that of others, is delimited by the frame. Distance between oneself and the other interactant is fixed and completely deprived of social connotations. No matter what social variables are at stake, videochat interaction occurs within a fixed frame, which seems to erode culture-bound proxemic patterns. In other words, regardless of social and cultural positions, videochat communication systems seem to erase all kind of differences and to be favoring a flat representation of social and cultural identities.

Management of space in videochat is problematic as it *stages* a spatial setting and a spatial distance. Participants can in fact provide a more or less distant image of themselves, for example displaying a very close shot or a medium shot to their discourse partner/s. This may be achieved through webcam positioning, which is

never neutral, but is an integral part of interaction in that it may represent an *intentional act* on the part of each participant.

Digital kinesics

Just as proxemic patterns are linked to culture and individuals, and any visual analysis needs to take such variables into account, so the same may be said about kinesics. Very few body movements have invariant meaning within or across cultures as body movements cannot be translated as directly as verbal behavior can (Kendon 1990, 2004). Apart from messages conveyed by movements, the additional question of intentionality needs to be taken into account. In verbal communication, there is a *deliberate* attempt to convey a message to a recipient. With non verbal behavior, the question of intentionality is less clear-cut, because some actions may well be defined as intentional (e.g. deictic gestures, such as pointing to objects and people in the physical world, Norris 2004: 28). Others are halfway between intentional and unintentional (e.g. iconic or metaphoric gestures, possessing a pictorial content and mimicking what is conveyed or hinted at verbally, Norris 2004: 28). Others are unintentional (e.g. postural behavior, Norris 2004: 24-27). However, these generalizations do not purport to represent stable relationships between intentionality and behavior. Any specific interaction may present different configurations of intentionality from those sketched here, as any interaction is culturally, socially and individually determined. However, the question of intentionality is seminal in any visual exploration of body movements and overall interactional patterns in communication.

Kinesics has traditionally focused attention on hands and head movements (Kendon 2004, Martinec 2004), which have been studied as "action" behaviors, which consist of discrete units of analysis, including "onset" and "offset" points (i.e. distinct beginning and ending of actions). Some studies focus on verbal and non verbal congruence, analyzing speech units, pauses and timing of body movements (Boomer & Dittman 1964), while social psychologists are more concerned with inner states revealed by non verbal behavior (cf. Ekman & Friesen 1969). However, multimodal analysis does not deal with what people are thinking, but with what people are *communicating* in interactions.

In the context of videochats, body movements cannot be completely ascertained by analysts, since hand movements, for example, may be semi-visible and, as mentioned above, partly fall outside the video frame. Other body movements, such as postural shifts, may even *totally* fall outside the video frame and, as such, be

missed both by the other discourse partner and subsequent observers (i.e. multimodal analysts). Video frames do not include the entire body, but usually frame the face, and optionally a part of the torso, creating a *frozen yet living* image of users. The *frozen yet living* definition attempts to capture the contradiction in the video frame which includes a frozen, fixed and partial representation of the body and at the same time presents a living representation of speech, facial expressions and body movements. The latter are typically instantiated by hands and head movements and torso postural shifts. The discourse partner's face is thus a metonym for the body realized by another face: in fact one's own.

Reciprocating gaze and impossibility of eye contact

Gaze is one of the most effective resources in interpreting and making sense of a discourse partner's attitude, stance and behavior. Early research literature on visual behavior was concerned with a series of questions on different levels of communication.

As a dependent variable, gaze has been used to measure stable individual and group differences, the regulation of the flow of conversation, and the search for feedback in interaction. As an independent variable, it has been shown to influence emotional responses and cognitive attributions (Ellsworth & Ludwig 1972). In the context of videochats, significant questions deal with gaze direction and the meanings attributed to the other's visual behavior.

Studies in psychology, interactional analysis and ethnography of communication have focused on gaze and its role in interactions and have provided insights into the different levels of analysis and stratification of gaze functions. Argyle et al.'s seminal paper (1973) has illustrated different gaze functions which permit a systematic analysis in keeping with different interactional conditions. In other words, interactions are context-dependent and though some aspects are more likely to occur in interaction (e.g. obtain feedback), others may be absent or mutually exclusive (e.g. the relationship between gaze and intimacy vs. inhibition of gaze to avoid undue intimacy). Any analysis concerned with gaze is thus heavily dependent on the context in which communication occurs and while some aspects may be ignored, others need to be carefully taken into account (Harrigan 2005).

To what extent is turn-taking accompanied by gaze in videochats? Sacks, Schegloff and Jefferson (1974) were the first to find no ideal overlap in conversation and identified *transition relevance place* (i.e. the end of turn construction unit in conversation analysis) as the crucial point at which floor changes are negotiated.

Goodwin (1980, 1981) and Kendon (1967, 1990) demonstrate that turn-taking behavior in conversation is mainly regulated through gaze. Kendon shows that *transition relevance places* are anticipated through gaze, while Goodwin argues that having listener gaze is so important for the turn-claim process that turn claimants will restart their utterances until the speaker's gaze assures them they *do* have the floor. Telephone and computer mediated communication present problems in turn-taking which resemble those described in blind and sighted interaction, since gaze is a fundamental facilitator (Everts 2004). Research literature shows that subjects use more turns when they experience more gaze (Kendon 1967) and experiments have been carried out to show effects of eye gaze on mediated group conversations (Vertegaal & Ding 2002).

While telephone conversations do not provide eye contact cues, videoconferencing systems (including videophone) give the incorrect impression that the remote discourse partner is avoiding eye contact. Some videoconference systems have cameras placed in the screens that reduce the amount of the difference in parallax (Otto et. Al 1993). This issue is also being addressed through research that generates a synthetic image with eye contact using stereo reconstruction or cutting-edge systems able to simulate eye contact between a three-dimensionally transmitted remote participant and a group of observers in a 3D teleconferencing system (Jones et Al. 2009).

Management of gaze is bound by many factors, such as context, culture, media and so on and, as such, reflections on its use cannot claim any kind of universality, but some uses of gaze may be intuitively and successfully interpreted, especially in conversations. Eye contact is in effect an almost universal feature of face-to-face interactions and, even though direct gaze in some cultures is permitted only in intimate contexts, exchanging looks, staring and eye contact are natural properties of spontaneous conversation. This feature is problematic in videochats and a fundamental question to be posed is how analysts can analyze gaze in spontaneous but technologically-mediated conversation. Eye contact is impossible in videochats, because participants *either* look at the frame where the discourse partner's image is projected *or* at the webcam, thus giving the other the illusion of direct gaze. However, in videochats virtual gaze is *never* reciprocal: if Participant 1 looks at the webcam, Participant 2 feels he/she is being looked at straight in the eye. But this perception cannot be reciprocal, since Participant 2 cannot reciprocate gaze *at the same time*.

Giving the other discourse partner the impression of direct gaze paradoxically excludes the possibility of seeing the other *tout court*. The problem is solved partially, only when the webcam is in the centre of the screen. However, not even

embedded webcams allow perfect eye contact. Simulating eye contact in virtual worlds is far from easy.

Vector analysis is thus complex, if not impossible, since the impossibility of eye contact alters traditional parameters of analysis. Gaze is thus still a facilitator for conversation in videochats, but its incidence and role is not easy to gauge. Should gaze be absent, as in the case of telephone conversations, other facilitators, such as prosodic elements, take over, but the peculiar presence of non direct and non reciprocal gaze requires rethinking of well-established models of visual analysis to be undertaken (van Leeuwen & Jewitt 2001; Baldry & Thibault 2006; O'Halloran 2004).

The current impossibility of eye contact in videochats implicitly underscores the virtual conversational nature of videochat.

Communicative apocalypse: visual or verbal decline?

Technological advances create new opportunities for interaction (Kress 2010), but a general increase in the potential for interaction (i.e. a wealth of new Web 2.0 texts, such as blogs, chats, forums, videochats etc.) does not necessarily equate with a proportional increase in communication. However, a potential apocalypse of the word is countered by the potential apocalypse of the image. In other words, we need to be cautious in making assumptions about the extent, degree and consequences of the development of web-based video interactions.

Participants do not ultimately interact with each other in videochats, but with the medium. The context of situation is both shared and *not* shared at the same time. Each participant is an active producer of illusions for the other, in orienting his/her gaze, in exploiting verbal and non verbal strategies to create a fake (or virtual) environment for the other, which is only partially real. Ultimately, the backbone of reality is constituted by interaction with a machine. Verbal and non verbal strategies add to a sense of reality while dissolving it at the same time. From these considerations, it follows that digital environments neither nullify verbal language nor invalidate the power of visual statements. Web-based interactions stand at the crucible of digital apocalypses, as they blend word and images and shatter them at the same time.

These considerations, briefly touched on here, raise questions for various disciplines, such as linguistics, semiotics, social sciences, psychology and neuroscience, to name but a few. Seen from another perspective, these questions point to problems that need to be addressed with the help of many scientific tools,

drawn from many disciplines, further blurring the epistemological boundaries between them.

References

Argyle, Michael, Roger Ingham, Florisse Alkema & Margaret McCallin. 1973. The different functions of gaze. *Semiotica*, 19-32.
Baldry, Anthony P. and Paul J. Thibault. 2006. *Multimodal Transcription and Text Analysis. A Multimodal Toolkit and Coursebook with Associated Online Course.* London and New York: Equinox.
Boomer, Donald S. & Allen T. Dittmann**.** 1964. Speech rate, filled pauses and body movement in interviews. *Journal of Nervous and Mental Diseases*. 139, 324-327.
Ekman, Paul & Wallace V. Friesen. 1969. The repertoire of nonverbal behavior: categories, origins, usage, and coding. *Semiotica*. 1, 49- 98.
Ellsworth, Phoebe C. & Linda M. Ludwig. 1972. Visual behavior in social interaction. *Journal of Communication*. 22(4). 375-403.
Emmison, Michael & Philip Smith. 2000. *Researching the Visual. Images, Objects, contexts and Interactions in Social and Cultural Enquiry*. London: Sage.
Everts, Elisa. 2004. Modalities of turn-taking in blind/sighted interaction: better to be seen and not heard? In Philip LeVine & Ronald Scollon (eds.). 2004. *Discourse and Technology: Multimodal Discourse* Analysis, 128-145. Washington D.C.: Georgetown University Press.
Goffman, Ervin. 1971. *Relations in Public: Microstudies of the Public Order*. London: Allen Lane, Penguin.
Goodwin, Charles. 1980. Restart, pauses, and the achievement of mutual gaze at turn-beginning. *Sociological Inquiry*. 50 (3-4). 272-302.
Goodwin, Charles. 1981. *Conversational Organization: Interaction Between Speakers and Hearers*. New York: Academic Press.
Guattari, Félix. 1989. *The Three Ecologies*. Trans. by Ian Pindar & Paul Sutton. London: Continuum.
Hall, Edward T. & Mildred Reed Hall. 1990. *Understanding Cultural Differences*. Yarmouth, Maine: Intercultural Press.
Hall, Edward T. 1966. *The Hidden Dimension*. New York: Doubleday.
Halliday, Michael A. K. 1978. *Language as Social Semiotic. The Social Interpretation of Language and Meaning*. London: Arnold.
Halliday, Michael A. K. 2002 [1987]. Spoken and written modes of meaning. In M.A.K. Halliday. *On Grammar* (ed. J. J. Webster), 323-351. London and New York: Continuum.
Halliday, Michael A. K., Mathiessen Christian. 2004. *An Introduction to Functional Grammar*. Third Edition, London: Arnold.
Harrigan, Jinni A. 2005. Proxemics, kinesics and gaze. In Jinni A. Harrigan, Robert Rosenthal & Klaus R. Scherer (eds.). *The New Handbook of Methods in Nonverbal Behavior Research*, 137-198. Oxford: Oxford University Press.
Heath, Christian & Paul Luff. 2000. *Technology-in-Action*. New York and Cambridge: Cambridge University Press.

Heller, Monica. 1988. *Codeswitching: Anthropological and Sociolinguistic Perspectives*. Berlin: Mouton de Gruyter.
Herring, Susan C. (ed.). 1996. *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*. *Pragmatics and Beyond* series. Amsterdam: John Benjamins.
Holmes, David. 2005. *Communication Theory. Media, Technology and Society*. London: Sage.
Jones, Andrew et Al. 2009. Achieving eye contact in a one-to-many 3D video teleconferencing system, http://arnetminer.org/viewpub.do?pid=1227884. Retrieved October 15, 2010.
Kendon, Adam. 1967. Some functions of gaze-direction in social interaction. *Acta Psychologica*. 26, 22-63.
Kendon, Adam. 1990. *Conducting Interaction. Patterns of Behavior in Focused Encounters*. Cambridge: Cambridge University Press.
Kendon, Adam. 2004. *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
Kress, Gunther. 2003. *Literacy in the New Media Age*. London & New York: Routledge.
Kress, Gunther. 2010. *Multimodality: A Social Semiotic Approach to Contemporary Communication*. London & New York: Routledge.
Lakoff, Robin Tolmach. 1982. Some of my favorite writers are literate: The mingling of oral and literate strategies in written communication" in D. Tannen, (ed.), *Spoken and Written Language: Exploring Orality and Literacy*. 239-260. Norwood, New Jersey: Ablex, Vol. IX, Series Advances in Discourse Processes.
Lofland, Lyn. 1973. *A World of Strangers*. New York: Basic Books.
Martinec, Radan. 2004. Gestures that co-occur with speech as a systematic resource: the realization of experiential meanings in indexes. *Social Semiotics*. 14 (2). 193-213.
McLuhan, Marshall & Quentin Fiore. 1967. *The Medium is the Massage: An Inventory of Effects*. London: Penguin.
McLuhan, Marshall. 1964. *Understanding Media: The Extensions of Man.* New York: McGraw Hill.
Norris, Sigrid. 2004. *Analyzing Multimodal Interaction. A Methodological Framework*. London and New York: Routledge.
O'Halloran, Kay L. (ed.) 2004. *Multimodal Discourse Analysis: Systemic Functional Perspectives*. London and New York: Continuum.
Ong, W. J. 1977. *Interfaces of the Word*. Ithaca: Cornell University Press.
Ong, W.J. 1982. *Orality and Literacy: The Technologizing of the Word*. London: Methuen.
Otto, Maximilian et Al. 1993. Teleconferencing eye contract using a virtual camera. Proceedings CHI '93 INTERACT '93 and CHI '93 conference companion on Human factors in computing systems. 109-110. Amsterdam.
Poster, Mark. 1995. *The Second Media Age*. Cambridge: Polity Press.
Postman, Neil. 1985. *Amusing Ourselves to Death: Public Discourse in the Age of Show Business*. New York: Penguin.
Rheingold, Howard. 1994. *The Virtual Community: Homesteading on the Electronic Frontier*. London: Secker & Warburg.
Sacks, Harvey, Emanuel A. Schegloff & Gail Jefferson. 1974. A simplest systematic for the organization of turn-taking for conversation. *Language*. 50, 696-735.

Scollon, Ronald. 1998. *Mediated Discourse as Social Interaction: An Ethnographic Study of News Discourse*. London: Longman.
Scollon, Ronald. 2001. *Mediated Discourse. The Nexus of Practice*. London & New York: Routledge.
Simmel, Georg. 1921 [1908]. Sociology of the senses: visual interaction, in Robert E. Park, and Edward Burgess (eds.). *Introduction to the Science of Sociology*. 356-61. Chicago: Chicago University Press.
Sindoni, Maria Grazia. 2010. Models of verbal and non-verbal interaction in Web 2.0 textuality. *Bérénice*. XVI(43), 85-93.
Tannen, Deborah (ed.). 1982. *Spoken and Written Language: Exploring Orality and Literacy*. Norwood, NJ: Ablex, Vol. IX, Series Advances in Discourse Processes.
Tannen, Deborah. 2007. *Talking Voices: Repetition, Dialogue and Imagery in Conversational Discourse*. 2nd Edition. New York: Cambridge University Press.
van Leeuwen, Theo & Carey Jewitt (eds.). 2001. *The Handbook of Visual Analysis*. London: Sage.
Vertegaal, Roel and Yaping Ding. 2002. Explaining effects of eye gaze on mediated group conversations: amount or synchronization? In Proceedings of the 2002 ACM conference on Computer supported cooperative work, 41-48.
Wesch, Michael. 2009. *The Machine is (Changing) Us: YouTube and the Politics of Authenticity*, http://www.youtube.com/watch?v=09gR6VPVrpw. Accessed 26 November 2010.