# Social signals: from theory to applications

**Isabella Poggi · Francesca D'Errico ·
Alessandro Vinciarelli**

**Abstract** The Special Issue Editorial introduces the
research milieu in which Social Signal Processing origi-
nates, by merging computer scientists and social scientists
and giving rise to this field in parallel with Human–Com-
puter Interaction, Affective Computing, and Embodied
Conversational Agents, all similarly characterized by high
interdisciplinarity, stress on multimodality of communica-
tion, and the continuous loop from theory to simulation and
application. Some frameworks of the cognitive and social
processes underlying social signals are identified as refer-
ence points (Theory of Mind and Intersubjectivity, mirror
neurons, and the ontogenesis and phylogenesis of commu-
nication), while three dichotomies (automatic vs. con-
trolled, individualistic vs. intersubjective, and meaning vs.
influence) are singled out as leads to navigate within the
theoretical and applicative studies presented in the Special
Issue.

**Keywords** Social signals · Theory of social signals ·
Applications

I. Poggi · F. D'Errico (✉) · A. Vinciarelli
Roma Tre University, Rome, Italy
e-mail: fderrico@uniroma3.it

I. Poggi
e-mail: poggi@uniroma3.it

A. Vinciarelli
e-mail: Vinciarelli@glasgow.ac.uk

## Signals, social signals, Social Signal Processing

Within the new domains developed in the last decades at
the interface between Computing and Social Sciences, a
new research area is emerging, also thanks to impulse in
computational social science and organization engineering
(Pentland 2007): the area of "Social Signal Processing"
(Vinciarelli et al. 2012). The idea is to apply Signal Pro-
cessing techniques—automatic analysis and understanding
of physical signals—to capture the social import of signals,
that is, the information they bear about social facts, and the
social influence they exert over individuals, groups, and
their interaction. The area follows an intuition by Curhan
and Pentland (2007), two engineers fascinated by works in
psychology by Ambady and Rosenthal (1993) showing that
future developments in social relations can be predicted on
the basis of "thin slices" of interactions: there is probably
something, in the unaware subtleties of people's behavior,
which can tell us how things are going. So, why not try to
capture these physical signals, typically beyond deliberate
production and aware perception, by electronic sensors,
and predict their outcomes through statistical processing?

Since 2009, the European Network of Excellence SSP-
Net[1] (Social Signal Processing Network) has been dis-
seminating this idea by elaborating theory, methods, and
techniques in Social Signal Processing. The new area
shares some research aspects with neighboring domains
like Human–Computer Interaction, Affective Computing,
and Embodied Conversational Agents:

1. it is a highly interdisciplinary domain that accepts
   contributions from many different fields—from neu-
   rosciences to computer science, from psychology to

---

[1] www.sspnet.eu.

philosophy of language, from Linguistics to Ethology—and makes use of diverse methods—conceptual analysis, experiments, surveys, simulations;

2. it stresses multimodality, that is, it credits great importance to detailed analysis of any kind of body signals and to their combinations;

3. it progresses through a loop between theory and simulation, creation of software tools, application, evaluation, and advancement in theory.

But beside such commonalities, what features specifically distinguish Social Signal Processing from other areas?

The goal of this Special Issue is to set the foundations of the new field by defining the notions of signal and social signal and by providing samples of research in this area. In the following, we overview some of the previous theoretical and empirical work to set the stage for the definitions provided and research presented in the papers of this issue.

## Cognitive and social processes underlying signals

Research in the last decades, emphasizing that much of our interaction encompasses bodily signals, has investigated *embodied communication* and the multimodality of communicative interaction. On one hand, scholars in Social and Computer Science relied on research findings and techniques for analyzing and classifying body behavior developed in the decades 1960–1980, like, for example, Ekman et al. (2002) Facial Action Coding System (FACS) and other studies on facial expressions and gestures (Morris 1997); on the other hand, they investigated the cognitive processes that rule the intertwining between words and gestures or other body signals (McNeill 1992, 2000; Kendon 2004; Kita 2003; Poggi 2007), assessing their link with the mechanisms of mirroring and joint action (Knoblich and Sebanz 2008) and simulating aspects of body communication in Virtual Agents (Wachsmuth et al. 2008). According to Allwood (2008), we have "embodied communication" whenever an "information sharing" between two biological or simply physical bodies (e.g., humans or machines) is actualized that is a *property* (or *aspect*) of physical or biological bodies, or that occurs thanks to a bodily process taking place *in their bodies*, or is *linked to* a biological or physical body. In a more strict sense, Poggi et al. (2010) consider some signals or parts or aspects of signals, like half-open eyelids that convey relaxation, as "embodied" morphemes of gaze, in that they stem from some physiological mechanism and become expressive or communicative signals through a process of metaphorization and ritualization.

But what are the cognitive and social processes that underlie signals and communication?

Major efforts to understand perception, memorization, and representation of signals as well as their relationship with social interaction come from the field of *Social Cognition*, a socio-cognitive approach that studies how people interpret and attribute meaning to one's own and others' behaviors. Following the pioneering works of Bartlett (1932) on *schemes*, that point at the representational nature of knowledge and the role of memory in reconstructing previously stored events to make memories coherent with reference schemes, also categories like status, role, human groups, are shown to be uncontrollable and unintentional at least in their default working processes, since they respond to *laws of cognitive economy* (Fiske and Taylor 1991). Perception, memorization, judgment, in absence of high motivation and opportunity of time and resources, follow an automatic path; categories and schemes influence information on the basis of accessibility in terms of past experience or primacy effect (Higgins and Rholes 1976), and to have a coherent description of a person, people organize impressions as a whole starting from a few first elements (Asch 1946). Finally, categorization—the clustering of different elements on the basis of one shared condition—simplifies social perception and social judgment by making external stimuli more accessible and triggering sets of information focused on particular objects, interrelated and organized in schemes (Fiske and Taylor 2008). Even the discriminative behavior toward outgroup members is generally triggered on the basis of the automatic activation of membership categories (Bargh 2006).

Novel views on social interaction and signal exchange in it came from heterogenous, sometimes contrasting, lines of research: TT, the *Theory of the Theory of Mind*, and the *Intersubjectivity* perspective. According to TT, interaction is made possible by people's capacity of representing—possibly through understanding of facial and bodily expressive signals—each other's mental states (beliefs, goals, emotions). Such capacity is implemented in brain structures, for example modules, as claimed by studies that suspect impairment of these structures in autism (Williams et al. 2001; Oberman et al. 2005), where the capacity for a Theory of Mind appears to be disrupted (Baron-Cohen 1991, 1995; Gallese 2006).

Opposed to this is the perspective of *intersubjectivity*, according to which social interaction is made possible since the very first days for a human child by an inborn capacity for "the sharing of experiences" between two humans; this includes the sharing of emotion, attention, and intention (Brinck 2001) and is seen as preceding, not following, the capacities for figuring out the other's standpoint. Hence, making up a theory of mind of the other is

not considered, in this view, as a precondition or an ancestor of the "shared mind" (Zlatev et al., 2008).

Research in neuroscience, investigating the neural underpinnings of social cognition, seems to demonstrate that processes involved in social perception and behavior, like perception of conspecifics, as well as memory and behavior concerning others, activate different neural systems than the perception of objects (Adolphs, 1999). Moreover, the discovery of mirror neurons—the neurons activated not only by one's motor action but also by the perception of action in a conspecific (Rizzolatti and Arbib 1998)—demonstrates how humans are programmed for empathy (Gallese 2006), the representation of self and others (Uddin et al. 2008), learning from others through imitation (Meltzoff and Decety 2003; Iacoboni 2005) and joint action (Vesper et al. 2010). On these bases, some scholars even argue that "interactive processes are more than a context for social cognition: they can complement and even replace individual mechanisms" (DeJaegher et al. 2010: p. 441).

In studying the ontogenesis and phylogenesis of communication, Tomasello (2008) posits that while the communicative gestures of apes are only acts of requesting, possibly implying practical reasoning and understanding of others' goals, a human's request also entails recursive mindreading, joint intention, and attention; this gives rise to, but at the same time is made possible by, true cooperation: in the context of mutual collaboration aimed at a common goal, developed for instance in group hunting, "the tendency to request collaboration and to supply help in return, as the initial motive of human cooperative communication, could flourish because in this context helping my partner helps me" (p. 240). Thus, from acts of request, the acts of informing develop, which reveal a capacity for mutual expectations of cooperation and communicative intentions; and at the end of the development process, information about one's attitudes is not only aimed at letting the other know of them, but at *sharing* them, which along with the motivation to *be like* others, and to *be liked by* them, leads to the development of norms and conventions, in the social as well as in the communicative domain.

## Lines of thought in the study of social signals and communication

Two parallel dichotomies underlie these studies: one between the individualist and the intersubjective standpoint, and one between automatic and aware-controlled processes, which has to do with the never-ending pendulum between stressing a highly self-aware view of cognition as opposed to more low-level mechanisms of behavior (like in the "cognition—emotion" diatribe that opposed Lazarus (1982) to Zajonc (1980)). This pendulum shifted research

during four decades from a dominant behaviorist perspective to a prevailing cognitivist one. However, the exaggerated stress on high level cognitive devices, privileging deliberative versus reactive decisions, did not leave room for irrationality or other less conscious mechanisms of behavior. Only later developments led to re-discover the importance of automatic and unconscious processes. Both dichotomies stem, in a sense, from the tendency to privilege—or even reduce to—a single feature or root for any given psychological device. But from the point of view of a theory of signals, these dichotomies, far from constituting an aporia, contribute to provide a wider and more articulated view of signals and social signals.

### Automatic versus controlled

On the automatic-controlled side, discovering the existence and importance of automatic processes leads one to accept that signals (see for instance pupil dilation as a signal of arousal) can be both produced by a Sender and received by a Receiver (that will be possibly influenced) without awareness, and this is often so for body signals. This is an important reason why signal processing techniques, based on bare statistical processing of physical measurements, may be effective in capturing signals exchanged during human–human communication that are not consciously perceived by human observers. Some of Pentland's (2008) "honest signals," acoustic and visual features detected during salary negotiations and speed dating encounters, are maybe difficult to spot for human observers, but cannot escape the automatic eyes of machines. The same principle applies to the development of artificial agents that generate signals like those that people use during social interactions (e.g., artificial smiles, synthetic facial expressions, intonation, etc.). If automatic processes actually regulate most of our behavior and interactions, artificial signals should produce the same effect as natural ones, as shown in the pioneering works by Nass and Brave (2005).

### Individualistic versus intersubjective

In the opposition between individualistic and intersubjective standpoints, as far as communication is concerned, the last decades have seen a shift from transmissive models of communication, like those of Shannon and Weaver (1949) or Jakobson (1953), where the burden of communication is borne only by the Sender, to the interactive or dialogic models of communication, in which meaning is negotiated or co-constructed (Ghiglione 1986) and information is shared (Allwood 2008) by the interactants.

Yet, while co-construction is a plausible account for the achievement of knowledge and for common planning and cooperation—when people discuss, for example, at the end

of the discussion, everyone has gained more knowledge thanks to others' ideas, and ideas themselves are built up together—this is not so plausible for communication: if communication is necessary to socially constructed knowledge, not necessarily is communication socially constructed.

In a sense, the extreme dialogical and constructivist views of discourse, according to which all meaning is negotiated together and none of the interlocutors owns one's meaning completely, do not take sufficiently into account the Sender's communicative intentions and goals and take its freedom and free will away, somehow revealing an underlying conviction on the individual's powerlessness and its complete determination by social context.

In a more individualistic view, first comes information that is owned by the individual (whether co-constructed thanks to previous communicative interaction with others, or simply conceived by oneself), and the goal of communicating it; then comes communication, by text, discourse, dialog, discussion, and this may finally give rise to new social construction of knowledge. Co-constructed knowledge just arises thanks to communication, but communication per se is initiated on the basis of an individual act of (conscious or unconscious) will: a goal of communicating. In this perspective, communication and social action are not always, or necessarily, or ineluctably determined by the environment; and to the extent to which they are, nonetheless an individual with one's mind, goals, and intelligence can anyway put a filter between one's behavior and the social environment, making oneself free to have goals of one's own. And the same freedom and flexibility generally holds for the reception of communicative signals by the Addressee: s/he can either accept the Sender's message and comply with what s/he says/wants or not.

Of course, all of this holds for the human part of signaling and communication, and, moreover, for the conscious part of it. This is where the two mentioned dichotomies, controlled versus automatic processing and individualistic versus intersubjective perspective, cross with each other.

In fact, the individual freedom of sending, interpreting, and being influenced by signals may be suspended for the sake of very important biological goals such as feeding or mating, or in the urgency of emotion expression, which has an important adaptive value. But apart from cases where communication makes an appeal to biological impulses, like in emotional persuasion (Miceli et al. 2006), communication may be seen as an expression of free will and reflective interpretation on the part of both Sender and Addressee.

Also from the point of view of social cognition in general, the focus on the primacy of intersubjectivity should be carefully discussed. That humans are social animals, as Aristotle pointed out, is far too trivial to bother arguing: the fact itself that a newborn human being cannot survive without another's care means that we are social

from the very first day of our life. Yet, while some scholars argue that "intersubjectivity can be regarded as a precondition for discourse itself to occur" (Nathan and Alibali 2011, p. 258), one might object that intersubjectivity itself cannot do without conscious or unconscious exchange of signals. Moreover, the intersubjective stand often presents too an optimistic cooperative view of social interaction and social relationships (Castelfranchi 1992). In real life, sometimes cooperation holds only at a surface level, the level necessary to share some rules of the game, but not at the level of playing it, that is often characterized by the opposition of individual goals. Of course, in arguing, quarreling or fighting people "cooperate," in that they stick to and exploit the rules of interaction; yet, they use them to pursue goals that are not shared, but are, to the opposite, in concurrence, in competition or in conflict.

To sum up, an internalistic account of social interaction might be (negatively) viewed as stemming "from the 'possessive individualist' cast of Western culture (and capitalism)" (Zlatev et al. 2008, p. 8); but in a completely intersubjectivistic account the individual might appear too other-determined, and his/her action never be completely imputed to him/her but subject to diffusion of responsibility and likely to be always justified by social context.

These two approaches, far from being totally incompatible, provide two different points of view from which different aspects of communication can be investigated: cognitive approaches mainly consider the Sender's intentions, intersubjective approaches consider meaning in its contextualized form.

Meaning versus no meaning

A third difference of positions concerning social interaction and signaling is in the analysis of signals and the attributions of meanings to them. Some scholars are very skeptical concerning the existence of systematic correspondences between specific signals and specific meanings and tend to credit signals and cues only with influence, as opposed to proper information. Partly this criticism is directed toward the use of analyzing animal signals with categories taken from the analysis of human languages; but this fight against what is called a "deterministic relationship between signals and states" (Brunet et al., 2012, this volume) cuts across disciplines, from Ethology (Mehu and Scherer, 2012, this volume) to Psychology and Philosophy of Language (Wittgenstein 1953), to the point of reducing the meaning to use, and seeing it as always floating according to context.

Other scholars, on the contrary, strongly maintain the existence of systematic correspondences between signals and meanings and actively work at the construction of lexicons of gestures and other body communication systems (Efron 1941; Morris 1997; Poggi 2007), or at detailed

analyses of single bodily communicative items (Müller 1998, 2004; Kendon 2004).

## Lines of research in social signals

The theoretical positions with opposite stress in the dichotomies above—individualist versus intersubjectivist, conscious versus automatic, meaning versus no meaning—are all represented in the papers of this Special Issue, while the empirical studies presented tackle examples of signaling in both acoustic and visual modalities, and at both the conscious and the automatic level.

For example, the pan-informational and pan-intentional views of Castelfranchi (2012, this volume) and of Poggi and D'Errico (2012, this volume), seeing signals as necessarily bearing information and as regulated by intentions, social ends, or biological functions impinging on the Sender, whether an individual or a collective system, contrast with both Mehu and Scherer's (2012, this volume) paper, for whom the evolution of signals depends on the influence they exert on the Receiver, and with Brunet et al.'s (2012, this volume) work, that denies a proper meaning to politeness signals, seeing them as fluctuating across contexts.

The first seven papers of this Special Issue are mainly focused on theoretical models of sociality and social signals.

Mehu and Scherer (2012, this volume), reminding the contribution of social psychology, ethology, and evolutionary biology, and stressing the importance of selective pressures in the evolution of social cognition and signaling, distinguish the notions of informative cue and communicative signal. After surveying the mechanisms of production and interpretation of social signals and cues on the basis of research in Social Psychology concerning interpersonal relationships and emotional expression, and proposing a model that integrates them with contextual information, they suggest that computer-based processing systems should be modeled more in terms of functional significance than in terms of absolute conceptual meaning. Their emphasis on the "influence" as opposed to the "information" side of signals, which they tend to pass over, taking the notion of information as almost necessarily abstract and conceptual, contrasts with Castelfranchi's (2012, this volume) paper.

Castelfranchi (2012, this volume) focuses on the notion of "mind ascription": he explains how people attribute mental representations to other people even without awareness, in a fast and automatic way and without complex reasoning, but simply on the basis of mental structures like scripts and roles, and through heuristics, categories, and prejudices, or even by default. He argues that social interaction necessarily requires assumptions, even if sometimes unconscious and implicit, about the other's beliefs and goals, which are necessarily assumed to be there, allowing us to act "as if" they were so.

The same model based on the notions of goal and belief is adopted by Poggi and D'Errico (2012, this volume), who define as social signals those communicative or informative signals or cues that directly or indirectly provide information about "social facts": social interactions, social emotions, social attitudes, evaluations and stances, social relations, and social identities. Sticking to a firm conviction on the possibility of attributing specific meanings to social signals, they consider the role of context in their processes of production and interpretation; viewing context as the set of beliefs provided by current perception and previous knowledge about Topic, Sender, Addressee, and their roles and relationships, they describe how information coming from signals and their context can mix up, not only to change the output meaning of a signal but also to frame it into a clear framework that accounts for it and adds more information.

Brunet et al., (2012, this volume) contrast two models of politeness: one by Brown and Levinson, who in their Goffmanian account view politeness in terms of protection of a social interactant's face, and one by Arndt and Janney, who distinguish a conventional politeness, defined as respecting social rules and etiquette, and an interpersonal politeness, seen as consideration of other people and their feelings. They propose a view of politeness as a communicative state, more precisely a stance that may be expressed through social signals, which may give rise, though, to incorrect inferences. While presenting their framework of social signals, the authors provide examples of politeness signals that exploit different expressions across media (e.g., etiquette in face-to-face communication, as opposed to netiquette in computer-mediated communication)—and across situations—for example, people using more signals of politeness when they have to evaluate others than when they share personal experience. They argue that any signal can be interpreted only by taking into account context, which in turn encompasses personality and culture, and point out that even "small deviations from a pattern that would signal politeness can create a pattern that signals the exact opposite."

Cipriani and Del Re (2012, this volume) introduce Visual Sociology, a research area aimed at taking up as data for sociological research the visual elements—nonverbal actions, conversations, smiles—that are a crucial part of the symbolic events traditionally studied by Sociologists: rites, liturgies, power and legitimation structures, affiliations. Visual sociology is seen as a kind of study that can validly integrate sociological research by encompassing the richness of information offered by multimodal social signals. The authors provide and overview of several methodological issues in this area, including the problem of how to attribute a meaning to a pose or action and how to

integrate raw information given by the visual data with other sources of knowledge but also emphasize the perspective shifts induced by the widespread use of visual documents, like the subject becoming the object of research—as is the case in videobiographies. They also emphasize the fruitful link between visual sociology and qualitative research, witnessed by the numerous software packages for automatic qualitative analysis that allow taking into account not only verbal data but also images, pictures, videos, iconic material in general.

Giardini (2012, this volume) deals with gossip, a socially complex behavior through which people gather information about an absent third party, while testing their own competence and reliability, thus reaffirming their membership in a group. Through gossip people acquire information about their peers, while creating and revising social bonds, promoting cohesion and norm compliance. The risk for unreliability of gossip spread by people who cheat is mostly avoided by two mechanisms: deterrence—the fear of being punished by the reputation of cheating—and transmission—the fact that in gossip, the source of information is anonymous, and this reduces the consequences of cheating. This allowed the evolutionary stability of gossip, making it a reliable signal on which human societies developed.

Leone's (2012, this volume) paper deals with social signals in educational interaction. By relying on Vygotskij's notions of scaffolding and of zone of proximal development, she deals with the teachers' and mothers' signals aimed at regulating the scaffolding processes, those by which a more knowledgeable person, to foster the learning of new notions or competences of a less knowledgeable one, starts helping him in his first trials but then lowers the level of help, trusting his "zone of proximal development," and thus fostering his autonomous learning. In observing primary school teachers interacting with their pupils, and mothers with their children, the importance of grasping such signals becomes clear for example when a student is already able to perform some action but the teacher cannot tell this from his behavior and over-helps him, with the effect of blocking his activity and preventing subsequent learning.

The next five papers present experimental studies on social signals in different modalities: written (Novielli et al. 2012, this volume), facial (De Kok and Heylen (2012, this volume); Van der Maaten and Hendricks (2012, this volume); and Ochs et al. 2012, this volume), and acoustic (Pesarin et al. 2012, this volume).

Novielli et al. (2012, this volume), within a research program aimed at developing an intelligent virtual agent that applies natural argumentation techniques to persuade the users to improve their eating habits, study written or oral verbal interaction of Users with an Embodied Conversational Agent playing the role of artificial therapist in the healthy eating domain and compare their reactions to persuasive attempts. In two Wizard of Oz experiments that vary Users' different level of motivation, the authors find out, quite unexpectedly, that those who volunteered to interact with the Agent use fewer signals of emotion, irony, small talk, than other Users.

De Kok and Heylen's (2012, this volume) paper tackles the notion of context with an empirical study on backchannel behavior. They first collected MultiLis, a corpus of interactions video-recorded in an experimental setup in which three listeners listen to a speaker, but the speaker only sees one. By comparing backchannel behaviors of the "concealed" and "displayed" listeners to the same speaker, they investigate the notion of "response opportunity" and the dependence of the response on characteristics of the speaker's contribution. A graded optionality in response opportunities is found, since in some cases all three listeners respond, and in others only a subset of them. Though some differences result from subjective measures (the displayed listeners report higher rapport than the concealed listener), objective data do not highlight differences between the responses of the "displayed" and "concealed" listeners as regards co-occurrence with the speaker's pauses or gaze, while the specific types of backchannel gesture displayed differ across listeners.

Van der Maaten and Hendricks (2012, this volume), in view of the automatic recognition of facial expressions exploiting computer vision and machine learning techniques, develop a system that automatically recognizes the facial Action Units as defined by Ekman and Friesen's FACS (Facial Action Coding System). By using the active appearance model, a sophisticated deformable template to model the appearance of faces, the system identifies the location of facial feature points, and from the face, it extracts features that point to specific Action Units, by detecting their presence through a time series classification model, the linear-chain conditional random field. The system evaluation through a large data set of videos with posed and natural facial expressions shows an agreement higher than 90 % between the system and human FACS labelers.

Ochs et al. (2012, this volume) created a web application, the *E-smiles-creator*, to enable users to create different smiles on the virtual agent Greta that considers several parameters of smile: amplitude, mouth opening, symmetry of the lip corners, lip press, cheek raising, duration, and velocity of onset and offset of the smile. Based on the corpus of descriptions of smiles of amusement, embarrassment, and politeness collected through this tool, they propose a decision tree learning algorithm to identify the different morphological and dynamic characteristics of these smiles, resulting in an algorithm that allows the agent to display various polite, embarrassed, or amused smiles depending on the context. An evaluation

study demonstrates the appropriateness of generated smiles in different scenarios.

Pesarin et al. (2012, this volume) present a semi-automatic approach to the detection of conflict in conversation on an acoustic basis. They start from the assumptions of Conversation Analysis that turn organization (who speaks when and how much) provides information about social aspects of a conversation and that a higher frequency of turn overlapping, interruptions, and speaker changes can be a cue to conflict. After manual segmentation of debates into conflictual and non-conflictual intervals, they present the structure, and an evaluation experiment, of a conflict detection process in four modules, encompassing (1) the extraction of steady conversational periods (the time intervals during which the configuration of the conversation is stable), (2) a machine learning generative model that through Markov chains training finally provides a description of conversation dynamics, (3) a generative score space to distinguish configurations based on the more discriminative parameters, (4) segmentation that classifies intervals as conflictual or non-conflictual.

Esposito and Esposito, tackling affective communication, a relevant type of social signals, survey some studies on the automatic recognition of emotions from voice. Though not providing specific rules on how to implement intelligent emotional interfaces, and acknowledging that the studies reported give contradictory results, they take them as a cue to anticipate requirements for research in the field, and hints for a theoretical framework to extract rules from multimodal emotional data. They argue in favor of a holistic approach in order to the computational handling of affective states in speech that should encompass attention to the signal, the feature processing, and the computational model at the same time, and that only taking context into account might achieve a more accurate recognition of emotion in speech.

# References

Adolphs R (1999) Social cognition and the human brain. Trends Cogn Sci 3(12):469–479

Allwood J (2008) Dimensions of embodied communication—towards a typology of embodied communication. In: Wachsmuth I, Lenzen M, Knoblich G (eds) Embodied communication in humans and machines. Oxford University Press, Oxford, pp 257–284

Ambady N, Rosenthal R (1993) Half a minute: predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. J Personal Soc Psychol 64:431–441

Asch SE (1946) Forming impressions of personality. J Abnormal Soc Psychol 41:258–290

Bargh JA (2006) Social psychology and the unconscious: The automaticity of the higher mental processes. Psychology Press, Philadelphia

Baron-Cohen S (1991) Precursors to a theory of mind: understanding attention in others. In: Whiten A (ed) Natural theories of mind: evolution, development and simulation of everyday mindreading. Basil Blackwell, Oxford

Baron-Cohen S (1995) Mindblindness: an essay on autism and theory of mind. MIT Press, Cambridge

Bartlett FC (1932) Remembering: a study in experimental and social psychology. University Press, Cambridge

Brinck I (2001) Attention and the evolution of intentional communication. Pragmat Cogn 9(2):255–272

Brunet P, Cowie R, Donnan H, Douglas-Cowie E (2012, this volume) Politeness and social signals. Cogn Process 1–7. doi:10.1007/s10339-011-0418-8

Castelfranchi C (2012, this volume) Ascribing minds. Cogn Process. doi:10.1007/s10339-011- 0423-y

Castelfranchi C (1992) No more cooperation, please! in search of the social structure of verbal interaction. Communication from an artificial intelligence perspective. NATO ASI Ser 100:205–227

Cipriani R, Del Re E (2012, this volume) Imagination and society: the role of visual sociology. Cogn Process. doi:10.1007/s10339-012-0433-4

Curhan J, Pentland A (2007) Thin slices of negotiation: predicting outcomes from conversational dynamics within the first 5 minutes. J Appl Psychol 92(3):802–811

De Jaegher H, Di Paolo EA, Gallagher S (2010) Can social interaction constitute social cognition? Trends Cogn Sci 14(10):441–447

de Kok I, Heylen D (2012, this volume) Analyzing nonverbal listener responses using parallel recordings of multiple listeners. Cogn Process 1–8. doi:10.1007/s10339-012-0434-3

Efron D (1941) Gesture and environment. King's Crown Press, New York

Ekman P, Friesen W, Hager J (2002) Facial action coding system (FACS): manual. A Human Face, Salt Lake City

Fiske ST, Taylor SE (1991) Social cognition, 2nd edn. McGraw Hill, New York

Fiske ST, Taylor SE (2008) Social cognition: from brains to culture. McGraw-Hill, New York

Gallese V (2006) Intentional attunement: a neurophysiological perspective on social cognition and its disruption in autism. Brain Res 1079:15–24

Ghiglione R (1986) L'homme communiquant. Colin, Paris

Giardini F (2012, this volume) Deterrence and transmission as mechanisms ensuring reliability of gossip. Cogn Process. doi:10.1007/s10339-011-0421-0

Higgins ET, Rholes WS (1976) Impression formation and role fulfillment: a "holistic reference" approach. J Exp Soc Psychol 12:422–435

Iacoboni M (2005) Neural mechanisms of imitation. Curr Opin Neurobiol 15:632–637

Jakobson R (1953) Essais de Linguistique généra1. Minuit, Paris

Kendon A (2004) Gesture. Visible action as utterance. Cambridge University Press, Cambridge

Kita S (ed) (2003) Pointing: where language, culture, and cognition meet. Lawrence Erlbaum, Mahwah

Knoblich G, Sebanz N (2008) Evolving intentions for social interaction: from entrainment to joint action. Philosoph Trans R Soc Lond B Biol Sci 12; 363(1499):2021–31

Lazarus RS (1982) Thoughts on the relations between emotion and cognition. Am Psychol 37:1019–1024

Leone G (2012, this volume) Observing social signals in scaffolding interactions: how to detect when a helping intention risks falling short. Cogn Process doi:10.1007/s10339-011-0422-z

McNeill D (1992) Hand and mind. University of Chicago Press, Chicago

McNeill D (ed) (2000) Language and gesture. Cambridge University Press, Cambridge

Mehu M, Scherer KR (2012, this volume) A psycho-ethological approach to Social Signal Processing. Cogn Process

Meltzoff AN, Decety J (2003) What imitation tells us about social cognition. A rapprochement between developmental psychology and cognitive neuroscience. Philosoph Trans R Soc Lond 358:491–500

Miceli M, Poggi, I, de Rosis F (2006) Emotional and non-emotional persuasion. IJIS, Special issue on Natural Argumentation

Morris D (1997) Manwatching. A field-guide to human behavior. Jonathan Cape, London

Müller C (1998) Redebegleitende Gesten. Kulturgeschichte—Theorie—Sprachvergleich. Berlin-Verlag Arno Spitz, Berlin

Müller C (2004) Forms and uses of the palm up open hand: a case of a gesture family? In: Müller C, Posner R (eds) The semantics and pragmatics of everyday gestures. Weidler, Berlin

Nass C, Brave S (2005) Wired for speech: how voice activates and advances the human-computer relationship. The MIT Press, Cambridge

Nathan MJ, Alibali MW (2011) How gesture use enables intersubjectivity in the classroom. In: Stam G, Ishino M (eds) Integrating gestures. John Benjamins, Amsterdam, pp 257–266

Novielli N, Mazzotta I, De Carolis B, and Pizzutilo S (2012, this volume) Analysing user's reactions in advice-giving dialogues with a socially intelligent eca. Cogn Process. doi:10.1007/s10339-011-0420-1

Oberman LM, Hubbard EM, McCleery JP, Altschuler EL, Ramachandran VS, Pineda JA (2005) EEG evidence for mirror neuron dysfunction in autism spectrum disorders. Brain Res Cogn Brain Res 24:190–198

Ochs M, Niewiadomski R, Brunet P, and Pelachaud C (2012, this volume) Smiling virtual agent in social context. Cogn Process. doi:10.1007/s10339-011-0424-x

Pentland A (2007) Social Signal Processing. IEEE Sig Process Magaz 24(4):108–111

Pentland A (2008) Honest signals. How do they shape our world. MIT Press, Cambridge

Pesarin A, Cristani M, Murino V, and Vinciarelli A (2012, this volume) Conversation analysis at work: de- tection of conflict in competitive discussions through semi-automatic turn-organization analysis. Cogn Process. doi:10.1007/s10339-011-0417-9

Poggi I (2007) Mind, hands, face and body. A goal and belief view of multimodal communication. Weidler, Berlin

Poggi I, D'Errico F, Spagnolo A (2010) The embodied morphemes of gaze. In: Kopp S, Wachsmuth I (eds) Gesture in embodied communication and human-computer interaction. Springer, LNAI 5934. doi:10.1007/978-3-642-12553-9_4

Poggi I, D'Errico F (2012, this volume) Social signals. A framework in terms of goals and beliefs. Cogn Process

Rizzolatti G, Arbib MA (1998) Language within our grasp. Trends Neurosci 21:188–194

Shannon CE, Weaver W (1949) The mathematical theory of communication. University of Illinois, Urbana

Tomasello M (2008) Origins of human communication. The MIT press, Cambridge

Uddin LQ, Davies MS, Scott AA, Zaidel E, Bookheimer SY et al (2008) Neural basis of self and other representation in autism: an fMRI study of self-face recognition. PLoS One 3(10):e3526. doi:10.1371/journal.pone.0003526

van der Maaten L, Hendriks E (2012, this volume) Action unit classification using active appearance models and conditional random fields. Cogn Process doi:10.1007/s10339-011- 0419-7

Vesper C, Butterfill S, Knöblich G, Sebanz N (2010) A minimal architecture for joint action. Neural Netw 23:998–1003

Vinciarelli A, Pantic M, D'Errico F, Heylen D, Pantic M, Pelachaud C, Poggi I, Schroeder M (2012) Bridging the gap between social animal and unsocial machine: a survey of Social Signal Processing. IEEE Trans Affect Comput 3(1):69–87

Wachsmuth I, Lenzen M. Knoblich G (2008) In: (eds) Embodied communication in humans and machines. Oxford University press, Oxford

Williams JH, Whiten A, Suddendorf T, Perrett DI (2001) Imitation, mirror neurons and autism. Neurosci Biobehav Rev 25:287–295

Wittgenstein L (1953) Philosophical investigations. In: Anscombe GEM, Rhees, R Basil (eds) Blackwell, Oxford

Zajonc R (1980) Feeling and thinking: preferences need no inferences. Am Psychol 35:151–175

Zlatev JT, Racine C, Sinha C, Itkonen E (2008) The shared mind: perspectives on intersubjectivity. Benjamins, Amsterdam