Conversational metacognition

Joëlle Proust

Institut Jean-Nicod, CNRS
ENS, EHESS, Paris

August 25, 2007

Keywords: metacognition, gesture, function, mind-reading, control

Introduction

My goal in the present chapter is to relate two fields of research that have been rarely – if ever – associated, namely embodied communication and metacognition. 'Embodied communication' refers to the process of conveying information to one or several interlocutors through speech and associated bodily gestures, or through gestures only. The term 'metacognition' was initially used to refer to the capacity of knowing one's own knowledge states, as a synonym for metamemory. There is no apriori reason, however, to restrict the object of metacognition to epistemic states. Therefore it was recently proposed that metacognition refers to all the psychological mechanisms allowing to evaluate and predict the *cognitive adequacy* of one's performances in processing information.[1] For example, one is able to evaluate the quality of one's percepts (metaperception), the impact of one's emotions on one's decisions (meta-emotion), or one's capacity to conduct reasoning or planning (metareasoning, metaplanning). In all these and similar cases, mental performances are monitored, either implicitly or explicitly, for their successes and failures.[2] On the basis of this ongoing monitoring, predictions can be reliably achieved, concerning success in a specific task at a specific moment. Metacognition is thus used to decide (e.g.) whether one can trust one's perception, or whether one is emotionally able to speak in public. Typically, one becomes conscious of the outcome of a given metacognitive evaluation through specific

[1] For lack of space, it is not possible to argue for this definition here. For a full defence, see Proust (2007).
[2] Evaluating oneself for cognitive adequacy entails both a retrodictive and a predictive relationship to performance: an evaluation has to be based on a comparison between a given performance and a stored norm, constructed through past success/failure ratio for that kind of performance.

embodied experiences, such as epistemic feelings, (a feeling of attending, of knowing, a tip of the tongue experience, an insight experience).[3]

Given this broad definition of metacognition, it is *prima facie* plausible that embodied communication crucially involves metacognitive interventions. Was my speech clear, coherent, was my gesture appropriate - did my pointing identify its intended referent? I propose to call 'conversational metacognition' the set of abilities that allow an embodied speaker to make available to others and to receive from them specific markers concerning his/her "conversing adequacy".[4] The hypothesis that will be explored in the present chapter is that embodied communication in humans involves metacognitive gestures. We will begin by exploring the common properties that they have. In order to do so, two kinds of objections will need to be addressed. The first is that what I called above 'conversational metacognition' might actually have nothing specifically metacognitive about it. The idea is that the kind of distributed control exercised in the course of a turn-taking exchange is entirely regulated by first-order, joint-action types of processes.

A second, alternative objection would insist, on the contrary, that metacognition conceived as a *procedural* form of self-evaluation does not have the resources allowing conversation to dynamically track, and adjust to, ever-changing multifarious conversational felicity conditions[5]. A metarepresentational capacity, as articulated in a full-blown 'Theory of mind', would in this view be needed to supplement a dynamic representation of joint action as well as a metacognitive capacity. These two objections will be addressed respectively in sections 2 and 3 below. We will show that joint-action regulation is not sufficient to allow embodied conversation to develop, and that theory of mind regulation is not necessary. A novel idea will emerge in the latter discussion: one of the important specific functions of metacognitive gestures might be to calibrate *the sense of effort* among participants. A final discussion will concern the respective roles of altruistic and Machiavellian pressures in conversational metacognition.

Let us summarize the chapter's aim: it does not consist in offering a new taxonomy, but rather in establishing the importance of studying a variety of gestures specializing in conversational metacognition - drawing on empirical research on gestures, on the psychology

[3] Koriat (2000). Some authors, however, propose that metacognition could occur without conscious awareness. See Reder (ed.) (1996).

[4] Conversational metacognition is grafted upon individual control-and-monitoring processes of ongoing conversation that have been independently studied (see Levelt, 1983). These processes are necessary to maintain the informational flow between communicators at an optimal level in terms of quality, quantity, and relevance.

[5] See Austin, 1962, 14 sq. Infelicities are « things that can be and go wrong » while uttering performatives.

of action, on general pragmatics, on social cognition and on the philosophy of biology. In order to do so, we first need to list the common properties of metacognitive conversational gestures, and to contrast them with other forms of gestures as well as with other forms of metacognitive engagements.

Section 1- Metacognitive gestures

Metacognition encompasses a set of procedures that allow cognitive systems equipped with it to predict or evaluate their ability to perform a given cognitive operation. These procedures allow the system to make a decision concerning the information currently used: is it adequate, does it need to be completed, revised, erased? Actually, prediction and retrodiction are closely associated: self-prediction relies on internal feedback (collected in similar past operations) to compare estimated output with a stored norm- the level of activity at which there is a good prospect of reaching the present goal. For example, one can estimate how much one knows about a given subject based on one's current feeling of knowing, a somatic marker that correlates with the stored norm.[6] Self-retrodiction relies on external feedback to compare observed output with the stored norm. For example, one may immediately realize that one's response to a given problem feels wrong. Note that in all these cases, metacognition seems to involve emotions and bodily inscribed feelings rather than abstract conceptual reasoning.[7]

In contrast with individual metacognitive feelings such as these – a kind of metacognition that can but does not need to be communicated -, conversational metacognition seems to involve specifically communicational control processes, which, for that very reason, have to be generally *distributed over several actors*. Thus, the feedback relevant to know whether an embodied utterance produced at *t* is satisfactory or needs repair can be gained both through an internal comparison process (as in other forms of metacognition) and through the on-line linguistic and embodied response(s) from the recipient(s). Explicit interrogations or other speech responses, but also facial movements, (in particular eyebrow movements and gaze orientation), head noddings, postures, hand gestures, rhythmic patterns in gesture sequences, inform the speaker of the cognitive adequacy of his/her intervention. Recipients show him/her their *degree of understanding* (from none to full), as well as the *emotional effect* of that understanding, (interest, surprise, boredom, disgust, overwhelmedness) and their *decisions* to

[6] See Koriat et al. (2006).
[7] See Proust (2007).

accept or reject the representational content (or the illocutionary act) conveyed.[8] An internal form of feedback however can *also* be used by the speaker to evaluate her productions. Talking about "internal" feedback should not lead one to think that such feedback constitutes a « private » store of information. It is generated by the social feedback gained in former conversational exchanges.[9]

A second distinctive property of conversational metacognition is that it provides a *multilevel* type of evaluation. As conversational analysis has shown, communicating agents need to keep track of the various task dimensions that are structuring talk in interaction. Thus an agent needs to monitor moment-by-moment *turn-taking* and *sequence organization.* (S)He must also keep track of his/her ability to refer and to achieve his/her illocutionary goals. An intriguing consequence of this multidimensional aspect of conversational metacognition is that a communicator needs to operate simultaneously on *different temporal frames* of gestural discourse, from the short-lived evaluation of his/her capacity to retrieve a proper name (while also keeping the floor) to the full-length appreciation of the success of a whole conversation.[10] A communicator needs to keep track of the specific sequence (s)he is in, and to permanently update her model of the exchange as a joint result of her embodied utterance and of the embodied responses and propositions that it prompted.

At this point, an objection needs to addressed. Why should we speak here of 'conversational metacognition' ? Why should not the distributed control exercised in the course of a turn-taking exchange be regulated by first-order, joint-action types of processes? The inner evaluation by an agent of the 'felicity' of an embodied communicative sequence would then just reflect the agent's monitoring of the relevant gestural and speech contents, by using the usual feedback for joint actions: rhythmic cues (indicating attunement with others' gestures and moods), verbal and gestural icons (evoking subgoals achievements and progress towards a common final goal), etc. This is an important objection, to which we will come back in section 2 at greater length. Let us start here with a clarification of the terms involved. How does one generally distinguish a cognitive from a metacognitive (mental) episode?[11]

---

[8] Are the recipients' responses always expressions of what we call "conversational metacognition" ? There is no simple answer to this question. As we shall see in section 2, it is arguable that a speaker can get metacognitive feedback from a recipient's reaction to what is said, that is not itself metacognitive. But recipients can set themselves in a metacognitive mode to appreciate an utterance in a reflexive way rather than simply react to it.

[9] Koriat (2006) for a discussion of the role of past monitoring and past control on present evaluation or prediction in judgments of learning. In the present article, it is speculated that the same global principles apply to conversational metacognition. Empirical studies, however, are yet to be performed.

[10] Although this has not been studied empirically in the context of conversation, the hierarchical analysis of action that has been shown in Koechlin et al.

[11] By a "mental episode", is meant any token of informational process, whether consciously executed or not.

Cognitive episodes are those whose function is to reduce uncertainty about states of affairs in the world. Metacognitive processes have the function to reduce uncertainty about one's own capacities. For example, predicting whether the next ball drawn from an urn will be red or black is a cognitive achievement. Predicting whether one will be able to solve a probabilistic reasoning task is a metacognitive achievement.[12]

If we now apply this distinction to conversational analysis, a gesture (or an utterance) is cognitive if its function is to refer to the conversational subject matter – an event in the world–, or to describe some property that it possesses, should or could, or will possess. A gesture (or an utterance) is metacognitive if its function is related to a speaker or a recipient evaluating how (s)he has been doing, or how well (s)he can hope to do, in the course of a given conversation in a given context. Examples of such metacognitive markers are offered by "Uhs" that allow a speaker to convey that (s)he will shortly be able to complete his/her utterance, by gazes and beats that indicate focused attention and motivation, and by various deictic gestures referring the audience back to a prior understanding that is now being taken as a common ground on which to elaborate further.[13] On the recipient's side, gestures such as "eye squinting" or "puzzled look" may reflect the intention to communicate one's own skepticism or difficulty in grasping the meaning of an utterance. One might again insist, that although these gestures have a specific function, they are generated by the same mechanisms that allow us to predict what others will do. We'll come back at length to this issue in the next section.

For now, let us take as a matter of definition that conversational metacognition has to do with checking one's (or another's) ability to convey an intended message through speech and gestures: it has to do with *"need repair questions"*: were the words and gestures produced adequate (intelligible, true, relevant?) Was I, was the speaker, in a position to make them? Was my (his/her) emotional expression congruent? Was my utterance accepted? It also has to do, less conspicuously, with *"should I" questions*: should I speak of X, given my poor

[12] For a more detailed explanation of this contrast, see Proust (2007).

[13] Our distinction between cognitive and metacognitive gestures is orthogonal to the tripartition between narrative, meta-narrative and paranarrative gestures offered by David Mc Neill. Metanarrative gestures refer "to the structure of the narration qua narration".(McNeill, 2005,) They include beats, which highlight referents, metaphorical gestures, which express various comments on the ongoing narrative, and spatializing gestures that contrast events or characters through their locations in space. (McNeill, 1992, 198-9). Another class, called "para-narrative", involves episodes in which a storyteller refers to her own experience and "adopts the role a participant in a socially defined situation of speaker and hearer" (McNeill, 1992) 199-200. This tripartition does not refer to the function of reducing uncertainty, and is also meant to account for descriptions rather than for general conversational needs (where promises, declaratives, expressives, also play their roles). Metacognitive comments can be expressed in words and gestures of the metanarrative and the paranarrative kinds.

memory?" "Should I admit that I did not understand what he just told?" It is important to note that these questions don't need to be consciously raised by the producer or by the recipient.[14] Note also that they only need to be raised in special circumstances (because some trouble is susceptible to arise, as appears on the basis of past experience compared with present or anticipated performance).

Another important observation has to be made at this point. Although we need to express the questions above in verbal terms in the present attempt at capturing how self-control occurs in conversation, they might not be necessarily couched in words, nor necessarily involve a conceptual representation of the communicational context. As we shall see in section 3, these questions are more likely to be raised and solved in a practical way rather by explicit conceptual reasoning. Indeed it has been observed that, in general, metacognitive processes are essentially procedural capacities designed to help us decide on how to act mentally.[15] Similarly, conversational metacognition might constitute a type of procedural self-knowledge designed to publicly control and monitor conversation moment by moment.

Function of gestures in embodied communication

This description however requires various specifications concerning the function(s) of metacognition. Clearly, some features of embodied communication may express metacognitive states *without having the function of expressing them*. To clarify this point, we need to briefly discuss the possible functions of gestures in conversation. According to the definition of *function*, for a conversational item to have a function, the item's presence in a conversation is explained by the fact that it typically produces an effect (has a meaning), and that it can be intentionally reproduced because it produces this effect (has this meaning).[16] Many features present in a conversation are not part of the meanings that their bearers intend to convey (they are natural indicators not symbols). For example, a pale face, tensions and rigidities in facial muscles or in posture may suggest that the speaker feels uncertain of being able to complete a turn or a discourse (think of certain painfully unprepared candidates to an oral exam). These embodied features – which are natural signs of fear, of anticipated failure –

[14] Empirical evidence concerning this point is still lacking in the case of conversational metacognition, a domain that has never been explored systematically. An indication in favor of non-conscious metacognition, however, is offered by research collected in Reder (ed.), (1996).
[15] See Smith (2003), Proust (2007) & Proust (in print). Section 3 will address the view that evaluation and prediction of capacity necessarily involves mind-reading abilities.
[16] On the definition of function, see Millikan (1993), and Proust (in print).

are not intended to communicate what they do.[17] The corresponding metacognitive states can be inferred by the audience, although they are not part of the speaker's utterance. The same holds for "adaptors", i.e. actions such as biting one's lips or grooming one's hair.[18] In contrast, a speaker may choose to intentionally express either in words or by gestures his/her current inability to offer a correct answer (for example, holding one's forehead in one's hands or scratching one's bent head, two metaphorical gestures for a searching mind). So we need to disentangle, in any piece of conversational analysis, the unintentional from the intentional gestures, natural signs from signs deliberately used to convey meaning. *A good primary indication for a gesture being intentional is that it can be controlled by the communicating agent*. In our previous example, the helpless student cannot control the amount of blood in his/her face, nor his/her muscular tensions. A gesture of shaking one's shoulders, as for releasing a burden, or extending the arms to the periphery displaying empty hands, on the other hand, are intentional ways of expressing inability or powerlessness. Invoking control, however, may help reject a number of embodied manifestations from being communicational gestures; it will not help *explain which exact function(s)* metacognitive gestures are serving.

Second, the expression we used above to state the function of conversational metacognition was deliberately vague: it "is related" to a speaker or recipient evaluating how (s)he has been doing, or how well (s)he can hope to do, in the course of a given conversation in a given context. *We need to determine which specific function conversational metacognition might play*. Research conducted on the function of conversational gestures in general will help understand the intricacy of the functions that could explain conversational metacognition.

It is currently widely accepted that conversational gestures in general contribute to communicating linguistic contents (although they may also be used to communicate independently from speech). This function is hearer-oriented: it is to help the *recipient* process the meaning of the spoken utterance, by presenting him/her with imagistic content –or to convey content without words. They do so both in a substantive way (by bringing additional information to the message content) and in a pragmatic way. The pragmatic contribution of gestures ranges from emphasizing structure and marking saliences, to indicating the illocutionary force or the interactional moves of the utterance (Kendon, 1995,

[17] The corresponding emotion, however, does have an independent, non communicative function (Griffiths, 1997). Some emotional expressions may be recruited as signals for communicational purposes, and used conventionally in meaningful gestures, such as joy or grief (controlled) facial expressions.
[18] See Ekman & Friesen (1972), Bavelas al, (1995).

Özyürek, 2002).[19] It is important to emphasize that a gesture token may well serve at once, substantive and pragmatic goals.[20]

Another intriguing hypothesis has been offered. Gestures might have primarily an executive or constructive role. (Krauss, 1998) They might aid speech production by facilitating lexical access (Hadar, 1989); or they might help the speaker to transform spatio-motoric thinking into analytic thinking (Goldin-Meadow et al., 1993, Alibali et al., 2000, McNeill, D. & Duncan, S. 2000). On this view, their main function is speaker-oriented: it is to help the *producer* construct his/her thought and, as a consequence, his/her conversational task in a multimodal way.[21]

At the present stage of the discussion of this debated issue, it is not obvious that we need to choose between these two options.[22] Susan Goldin Meadow and David McNeill (1999) have argued that the manual modality might specialize in mimetic representations (simulations of how things are), while the oral modality might be more adapted to conveying segmented combinatorial representations. This difference would not be radical, however. Some gestures (and some linguistic units) might exchange standard roles, with gestures coding abstract reference and speech mimicking natural or uttered sounds.

Additional reasons for combining communicative and constructive functions rather than selecting one against the other, are currently emerging in the literature. We will examine them in section 2 below and in the Appendix.

Functions of metacognitive gestures

Granting that there is no need to choose one function to the exclusion of the other, we can ask ourselves which functions are served by metacognitive gestures. Authors have studied some of them as "monitoring understanding" gestures (Clark & Krych, 2004) and illocutionary force markers (Kendon,1995, Bavelas & Gerwing, 2007). [23] What is metacognitive about them? If we take gestures and utterances to have primarily a communicative function (see above), we might propose that gestures and utterances are

[19] See in particular p. 247-8.
[20] See McNeill (1992) and Duncan (2006).
[21] This view predicts that speakers will use different gestures when used to reason about a scientific problem and to transmit scientific knowledge, which is actually found : Crowder (1996).
[22] See Jacobs & Garnham, (2007).
[23] Although "common ground gestures" (those gestures used to establish mutual understanding about space, reference, perspective, etc.) might be seen as belonging to metacognition, they do not need to have a metacognitive function: their specific aim is not to appreciate one's uncertainty, but to actively construct a shared world.

metacognitive if their function is to control and monitor *the epistemic capacities* of the receivers, by helping *them (rather than the producers)* to monitor the producer's current evaluation of his/her job at maintaining or repairing his/her contribution. Turning to the other option about function, embodied conversational metacognition might also have a producer-directed function, helping *him*/*her (rather than the recipient)* to represent for him/herself performance uncertainty, in order to keep salient the various associated pitfalls and take strategic measures to overcome them (like talking more carefully, more slowly, preparing interrupts or avoiding certain subject matters).

Now a further question is to understand why such functions had to emerge to make conversation at all possible. This is an issue of interest for philosophers who want to understand which constraints are shaping a given functional element, for cognitive scientists interested in the dynamics of mental activity and by researchers on speech gestures that aim to build up taxonomies. Given that these considerations are somewhat technical, we will present them in the appendix. Here is a short summary that will suffice for the present purpose. Efficiency of collaborative effort between several communicating individuals necessarily presupposes that:

1- there is a rhythmic pattern through which attentional processes can be jointly tuned (syntax, beats, postural sways);
2- basic rules can be learnt thanks to which collaborative effort can be minimized (such as Gricean maxims);
3- each participant is able to learn where s(he) stands relative to personal or inter-individual standards of informational adequacy.

These three constraints respectively shape:

1- the *communicational medium*, that needs to adjust to the attentional and computational capacities of the informational processing systems engaged;
2- the *complexity and flexibility* of the messages that can be conveyed (ie constraints on the communication goal);
3- the flexibility in *self-evaluation* (including a sensitivity in appreciating one's own success in effecting steps 1 and 2 and a capacity to revise occasional misfirings).

The three sets of constraints are embedded conditions. The first determines the dynamics for an exchange, the second states the conditions in which semantic content can be communicated; the third requires the ability to self-evaluate one's abilities to cope with the various conversational demands. If this analysis is correct, then we see why metacognitive

gestures need to emerge; they are part of the procedures that maintain an informational exchange on track; they don't deal with the basic establishment of the medium, nor with communicated content, but to how communication proceeds: it ensures stability in a changing world where knowledge is sparse and unevenly distributed.

Let us summarize. We have explored the conceptual possibility for metacognitive processes to develop in normal conversation through speech and gesture, with a function that can be recipient- as well as producer- oriented, having to do with the control of conversational (epistemic, motivational and social) adequacy. We saw that some of the gestures as studied in the literature do as a matter of fact serve these metacognitive functions.

2 - Joint action and the action-perception view on monitoring conversation

Now one might want to come back to the main objection against the proposal of singling out metacognitive gestures as a significant class of interactive gestures. On this proposal, as we saw, metacognitive gestures are needed to express self-directed uncertainty: to help recipients predict how the communication is going to develop, by making explicit both the producer's state of knowledge, or degree of involvement, and by checking with the recipients whether they grasp his/her intention. But, the objector might go, a simpler explanation of these gestures might be provided. In this alternative explanation, they merely contribute, along with other interactive gestures,[24] to the moment-by-moment monitoring of a dialogue; they structure the developing exchange by orienting turn taking; part of their role is to provide a rhythmic pattern, the other being to retrospectively check on the success of a turn. Just as gaze orientation can be exploited to monitor joint attention without being endowed with metacognitive significance, questions like "You know?" or "You see", directives such as "remember when P?" and the associated gestures, do not need to be given a metacognitive, self-directed interpretation. They are merely part of the basic control that allows a joint action to develop in a well-tuned and stable way. Similarly for the corresponding "back-channel" gestures: beats and looks, and referring gestures to listeners, are meant to elicit feedback to allow conversation to proceed. Gestures indeed have a remarkable advantage in this function over speech; they modulate the expression of illocutionary acts in ways that make them richer in meaning and more acceptable to the recipients. But the evaluative modulation they are

[24] See in particular Bavelas et al. (1995).

effecting is *not* metacognitive: it is part and parcel of the *control of the joint action* for constructing a social context. On this view, what confers a gesture a communicative function is not so much that "it has been produced with the intention that the recipient think that P as a result of his/her recognition of the producer's intention to get him/her to think so by producing this gesture" (along Gricean lines[25]). It is rather that it plays a causal role in a control loop distributed among several participants, and is produced and recognized because of this role.

This plausible objection might also invoke computational studies and neurophysiological evidence concerning action. The fundamental idea is that the brain simulates aspects of the sensorimotor loop in observing, planning and controlling actions.[26] The neural circuits involved in a specific action provide internal models for it. These models predict the sensory consequences of commands, whether on the physical world (own body and environment) or on the social world (others' behavior and associated mental states). Furthermore, these models can be activated by observing an action performed by another agent as well as by the actions performed by self. This allows agents engaged in a joint action to share closely similar versions of the action that they plan to perform. In conversation, as in any joint action, the cues that constitute the major dynamic steps organizing the internal model have to be publicly available to allow common alignment on feedback.

In summary, the evaluation by the participants of the « felicity» of an embodied communicative sequence would then merely reflect the normal competence that agents acquire when acting conversationally. This view has been articulated in the context of mirror-neuron contribution to conversation. The communicating agent A runs a simulation of his/her internal model of the next embodied utterance, and detects trouble before it actually occurs; similarly, listener B constructs a model of the developing utterance, and predicts trouble from various cues using his/her own experience.[27]

Crucial to this alternative perspective, is the view that self-prediction can be explained *without* invoking any form of self-questioning or self-evaluation. It is easy to reinterpret in these terms earlier findings by Schegloff (1984), that conversation control is built upon the

---

[25] See Grice, (1989).

[26] Three types of studies can be jointly used to make this point. 1) Ideomotor theories claim that observed actions are represented in the same format as executed actions (Prinz, 1997, Barsalou, 1999, Jeannerod, 1999). Neurophysiological evidence suggests that overlapping brain areas are activated during observation, action, planification and imagination of action, which is a main argument for a *simulatory* view of action representation. (Gallese et al., 1996, Decety et al., 1997). 3) computational studies invoke the need to *compare* observed and internal feedback to adjust command and to keep track of agency through efference copying. (Wohlpert et al., 2003)

[27] See Rizzolatti & Arbib, (1998), and Arbib (ed.), 2006.

notion of a projection space. Every expert communicator *knows how* to anticipate on the dynamics of a conversation: (s)he recognizes "what is being known and said before it has actually been done and said" (Schegloff, 1984, p. 268). A producer predicts that (s)he is going to have trouble ahead, and emits a "sound stretch" or produces 'uhs", or cutoffs both to warn the hearer and to reprocess during this time lapse the part of speech to repair. Reciprocally, the hearer knows how to decipher these error signals, and backtrack to the relevant part of the sequence where the repair occurs.

The most economical way of interpreting these capacities – in the alternative view - assumes that one and the same model is used both in language and in gesture for the hierarchical organization of sequences in conversation. This model is co-produced by the participants. They need to update it at each turn. If we consider a communicating system of several interacting agents as a joint action system, with a partly shared internal model of the task, standards of production that emerge from prior communicating experience, and various comparators, to identify and repair mismatches, projective goals encompass all kinds of goals. Repairs, self-serving metacognitive anticipations, social cues, are all *at the same level* because they are learnt through the very same type of conversational action.

Let us sum up. If ideomotor views can be extended to conversation, that is, if conversation is regulated by observed feedback, we should be able to identify in embodied communication, as Willem Levelt (1981) and Herbert Clark (Clark & Wilkes-Gibbs, 1986), among others, have done for linguistic discourse, control and monitoring devices in conversation *without caring for metacognitive abilities*. The reason for this parallelism is that there must be a level of control – joint action - that is common to speech and gesture ; dynamics of speech and of gesture are strikingly similar, and the ability to use verbal information or gestural imagery to convey various contents strongly suggests that they are intimately connected in their structure and in their realization.

Why the ideomotor view on conversation does not account for metacognitive gestures.

Several arguments however can be adduced against the proposed reduction of metacognitive to merely cognitive gestures. The first is that the ideomotor view on gesture may be only partly right, in the sense that some gestures – like emblems - can be learnt through an embodied simulation of perceived symbolizing actions while others can't. A reason to introduce this distinction is that, as Jacob & Jeannerod (2005) have shown, a simple ideomotor or resonance view is not sufficient to account for a communicative intention; for

example, it has trouble explaining how a conversational gesture such as a pointing-to-my-watch may acquire, given a context, either the meaning of "I want to leave the party", or the meaning of "my watch does not work". The same objection applies a fortiori to the gestures that allow agents to communicate about their own epistemic states. It may be that, for example, people learn the meaning of a "quick brow raising" as reinforcing stress on a word; but they have to distinguish this kind of brow raising from a metacognitive form of the gesture meaning that the utterance involves doubt and questioning (Ekman, 1979, Bavelas & Gerwing, 2007). True, as noted by Bavelas & Gerwing, speech intonation normally disambiguates this facial gesture's meaning. But the gesture can also be performed without speech, as in Jacob and Jeannerod's example. The point is that the relevant communicative intention can only be understood with the required flexibility if the recipient is able to simulate *himself/herself* as being in doubt about P, given a certain context of utterance. In other terms, the recipient must have (and apply) a dynamic model against which to evaluate what is uncertain in the present situation. Such a self-simulation involves much more than remembering the association between a facial gesture and a conversational outcome: it involves coupling an embodied state (observed in producer) with one's own epistemic state that P, through the possible mediation of another embodied state (in self), namely the somatic marker for that epistemic feeling (underlying "the sense of doubting") and a global model of the problem space where P is located.

The apparent force of the objection may be related to a difficulty in distinguishing general conversational control from conversational metacognition. Every form of action, individual or collective, needs to be evaluated against its own standard, as represented in a prior intention. Therefore, embodied communicators must compare not only the words, but also the gestures that they actually produce with those that they *intend* to produce and with *standards of production*.[28] They must therefore adjust their gesticulations to match the corresponding intentions and standards, and possibly correct them when necessary. This control function admittedly does not need to be metacognitive. Again, a gesture qualifies as metacognitive only if its content does not express a state of affairs in the world but rather is directed to one's own relative ability to perform some first order cognitive activity. Repairing speech or gesture does not, in general, qualify as metacognitive because it has a merely instrumental goal, namely to substitute an item to another, suppress ambiguity, provide

[28] See Levelt, (1983).

common grounding in reference, etc. Only those gestures expressing self-awareness in the informational or motivational conditions that affect performance in the task at hand do. Let us offer an example of gestural repair that does not count as metacognitive, but that does involve a comparator:

Example I - A 9 yr-old child involved in a ball game is quite animated and noisy in the school's playground. A teacher comes to order the group to stop shouting. The child crosses her arms, while protesting that the opponents were unfair, then quickly folds them behind her back.

Here the child corrects her gesture, moving from a defiant gesture directed at the other players to a submissive gesture directed at the teacher. The correction is prompted by the standards of gesture production in the school context, as opposed to the unqualified playground context. The gesture was inadequate, but was not made so by some metacognitive failure. It's rather a cognitive failure concerning the selection of a contextually appropriate gesture. In contrast, the following exchange elicits a metacognitive gesture from B:

Example 2

A : "Where did you put my tennis racket ?"

B: Frowning while looking up, then twisting the hands to the outside, thumbs up (no word uttered)

Here, the speaker recognizes in gestures that she cannot satisfy the conditions of felicity of the request for information. Her gestures however show, in addition, that 1) she is trying hard to remember, and 2) that her trying is nevertheless failing. This kind of self-simulation brings into play a piece of evidence that she is strictly speaking not requested to offer. She volunteers it to explain and justify why she does not abide by the request.

Although some interactive gestures may be explained in a crude first-order, observation/action way, many if not most conversational gestures need to be integrated within several differents systems of appraisal, some directly related to common goals, some to individual epistemic standards.[29]

A more general way of making the same point consists in using the distinction between three kinds of constraints mentioned in section 1 (and discussed in more detail in the Appendix). Efficiency of collaborative effort between several communicating individuals generally presupposes that three kinds of conditions are met, respectively shaping the

[29] As we saw above, these individual epistemic standards can be adjusted to the social context. More on this in the last section.

dynamics regulating common attention, the semantic content to be communicated, and the ability to self-evaluate one's abilities to cope with the various conversational demands.

The ideomotor approach to metacognitive gestures would only be promising if *the third* kind of constraints could be identified with *the first*, or at least with *a simplified account* for the second. If appreciating one's ability was a matter of observing rhythmic patterns and conforming to them, or a matter of simulating other's moves in order to grasp a motor intention, and thereby understand the content of a metaphoric gesture, or the referent of a deictic gesture, *then* we could indeed speculate that metacognitive gestures also have an intrinsic motor content. Simulating it would allow participants, in favorable conditions, to reach the same epistemic state as the producer's.

But the kind of simulation that is needed to perform metacognition in general belongs to *self-simulation*. Self-simulating can be illustrated by a basic directed recall attempt: you search your memory to retrieve a word, and derive from the simulated search (in comparison with previous similar attempts) predictions on your ability to retrieve the word. Motor activity is indeed taking place as an expression of mental search, but it is not followed by a change in the world: the activity involved is mental.[30] The only change in the world that needs to be monitored is the effect of the utterance on the recipient. So even though there are affective aspects in metacognitive gestures that afford "direct resonance", as in the failed remembering of our example (2) above, a recipient can only understand a metacognitive comment on a task if s(he) is able to perform the same *mental* task. These arguments allow us to conclude that conversational metacognition cannot be handled appropriately within an ideomotor, or a mirror-neuron framework.

Section 3 – A theory-of-mind view on conversational metacognition

An alternative objection, sketched in the introduction, reciprocally claims that metacognition conceived as a procedural form of self-evaluation cannot deal with the demands of conversation. A metarepresentational capacity, as articulated in a full-blown Theory of Mind, would in this view be needed to regulate both production and reception of gestural as well as spoken communication. This constitutes what we will call here "the mind-

[30] See Proust (2007) and (in print).

reading objection" to the very possibility of a conversational metacognition. Let us examine this objection in more detail.

Given the complex inferences that need to be performed to grasp the communicative intentions in most utterances, whether through speech or gesture, many theorists have speculated that only a subject equipped with mind-reading abilities would be capable of making sense, for example, of indirect speech acts. To grasp the communicative intention prompting the sentence "do you have salt"? (either as a genuine question, or as a request for salt), a speaker needs go beyond what is said to what is meant, by using something like Grice's cooperative principle or Sperber & Wilson's relevance theory. For some theorists, this process involves interpreting other's speech or behavior in terms of beliefs, desires and practical inferences, and having a way to select the most likely intention given the context. Interpreting mental communicated contents, on this view, entails metarepresenting the thoughts that the other conveys by speech or gesture. By metarepresentation, is meant a representation whose content includes: 1) a first-order representation, such as "I have the salt" and 2) the representation of an epistemic or a conative attitude directed at that content, such as, "he wants to know whether I have the salt" or "he desires me to pass him the salt". On other words, you cannot understand properly speech if you don't have the capacity to apply concepts such as "believing" or "desiring" to first-order contents.

This said, we can thus rephrase the mind-reading objection in the following way:

> *Such metacognitive gestures as eyebrow raisings or frownings, puzzled looks, metacognitive pointings, etc. can only be used and understood as the outcome of mental reasoning sequences, through metarepresentations containing the relevant concepts.*

Let us take, for example a certain squinted-eyes gesture, with the intended meaning

(1) [I know very well when P, and here it is not clear to me that P].

Let us suppose that this meaning is conventionally coded; the recipient B needs to apply mental concepts (knowledge, doubtfulness, etc.) to fully grasp the gesture's meaning. S(he) must grasp the conceptual content (1) as what is intentionally expressed by the facial gesture and identify the correct portion of A's or B's speech to which it refers. If we now suppose that some facial gestures express content by way of inferences rather by conventional coding, the recipient needs to reflect in the Gricean counterfactual way:

> "*Normally people only squint their eyes when they cannot see well. Currently, there is nothing to look at. No communicative gesture is made without a reason. The reason must be that the speaker wants me to recognize, by producing this gesture, her intention to express that s(he) does not see what is meant, when normally something should be made visible."*

In both cases (conventional or inferential), metacognitive communication is taken to depend on sophisticated inferences about others' states of mind in practical reasoning, that is: on metarepresentational capacities as represented in folk-psychology. This assumption, however, is confronted with notorious difficulties. The first problem is that children seem to be able to grasp utterance meaning, both in speech and in gesture, well before they master a theory of mind. One of the most important communicational gestures, declarative pointing, appears around 9 months, and by 18 months is used as a way of initiating joint attention acts with a social partner (Carpenter et al., 1998, Franco, 2005). Although joint attention can be described in rich mentalizing terms, as the realization that another person can be made to acquire new perceptions and beliefs on the world through a specific indicative gesture, early mastery of joint attention suggests that this capacity is rather controlled by an innate mechanism working as a precondition for theory of mind. Parallel studies in verbal and mind-reading development show that children learn to metarepresent with mental verbs after having mastered communication verbs (de Villiers, 2000, Harris et al., 2005). Developmental and clinical evidence (gathered from deaf children) suggests that conversational activity might be one of the driving forces in theory of mind acquisition, although theory of mind might help children refine their pragmatic expertise.

A second argument is that mental reasoning, were it necessary to evaluate the relative relevance of several interpretations of a speaker's intention, would require considerable memory resources and processing time. Inferring a speaker's meaning would, in these conditions, be too demanding for an individual to come up with the correct solution. Dan Sperber and Deirdre Wilson (Sperber & Wilson, 2002) have taken seriously this objection as well as the former one, and concluded that the procedure through which one infers a speaker's meaning "is not individually discovered, but is biologically evolved. It is an evolved module." On this view, mind-reading would encompass many different submodules documented by developmental psychologists (Baron-Cohen, 1995). An Eye Direction Detector exploits the correlation between direction of gaze and visual perception to attune one's perceptual attention to others'. An intention detection module interprets goal-oriented behavior as the intention to obtain a certain outcome. A Shared Attentional Mechanism allows human

communicators to perform declarative pointings with adequate joint-attention monitoring. Sperber & Wilson propose that an additional submodule recognizes communicative intentions. "Ostensive-inferential" gestures don't need elaborate mind-reading inferences to be produced or understood. The recipient merely takes the most economical coherent interpretation for the gesture, that is: the most *relevant*. The eye-squinting gesture, for example, involves two types of processing, following Sperber & Wilson's (1986) analysis of ostension:

> "First there is the information that has been, so to speak, pointed out; second, the information that the first layer of information has been intentionally pointed out" (50).

So if we come back to the embodied utterance above with the content (1), a recipient may understand what it means because (i) s(he) presumes that there is some interpretation of the utterance that is "the most relevant compatible with the speaker's abilities and preferences, and at least relevant enough to be worth the hearer's [/recipients] attention". (ii) s(he) follows a path of least effort in computing the cognitive effects of the gesture. iii) s(he) stops when expectations of relevance are satisfied.(Sperber & Wilson, 2002). Step (i) is not problematic; the "guarantee of relevance" forms the background needed for every communicational episode. It is established through prior experience that an ostensive-inferential behavior is meant to communicate something of interest to him/her.

A crucial element in S & W's solution is (ii): there must be an ordered sequence in which alternative interpretations come to mind, which is common to the producer and to the recipient. This sequence is what prompts a differential feeling of effort for the various portions of the sequence: an immediate inference does not cost much, whereas an inference where many steps have to be performed is perceived as more effortful. The theory says that the communicators don't need to explicitly think and compare different interpretations. They only need to make the necessary inferences in the same order and to have the same sense of satisfaction when reaching a given conclusion. But a new problem surfaces: how can one detect the differential amount of subjective effort associated with given computational demands? How can one store the "norm" for the kind of effort correlating with the correct solution?

The feeling of effort, actually, is a basic concept in theorizing about metacognition. We shall see below that the gist of S & W's view can be captured using a metacognitive, control-based semantic framework rather than a theory-of-mind approach to conversation. Our strategy for addressing the mind-reading objection in its S & W's revised formulation is to defend a deflationary approach to conversational understanding in general, and of

conversational metacognition in particular. The basic differences between this deflationary approach and S & W's submodular theory can be summarized in three claims.

1) The concept of a *communicative intention* can be understood implicitly in metacognitive terms – i.e., in procedural terms - or explicitly - in attributive metarepresentational terms.

2) Metacognitive development is *phylogenetically and ontogenetically* distinct from the development of metarepresentation and mind-reading.

3) Metacognitive capacities are task-specific rather than domain-specific

We will briefly examine these three claims, restricting our comments to aspects relevant to conversational metacognition.

1) The concept of a *communicative intention* can be understood implicitly in metacognitive terms – in procedural terms - or explicitly - in attributive terms.

S & W (2002) propose that intentions can be recognized by combining various salient cues and automatic associations, in a non-conceptual, modular way.[31] In other words, a communicator can correctly identify the intention to have the recipient recognize P as message, without using a full blown Gricean attribution of intention. A metacognitive approach uses a similar strategy.[32] Multimodal cues help recognize that a given movement has a communicative rather than an instrumental function. The cues however are not selected as a result of an innate processing bias; they are used because they have a specific functional status: they have been parsed and stored as feedback elements in prior cooperative exchanges. They now form building blocks for dynamical forward models of ongoing conversations. Some of the stored feedback elements are properties of "the world" (like exaggerated movements), and can thus be simulated at the mirror-neuron level. Others are epistemic feelings and somatic markers that correlate with dynamic properties of the informational flow. They are associated with hesitations, memory failures, tryings, etc. We saw in section 1 that such communicative events, when observed in another communicator, need to be self-

[31] For example, an *exaggerated movement* automatically captures others' attention. A movement performed *outside its instrumental context* also makes it open to a communicative interpretation. Being performed with an associated *gaze at the recipient* is a third cue leading to the proper interpretation.

[32] A strategy that does not need to posit an innately modular structure of mentalizing abilities. See Samuels (1998) and Proust (in print) for some of the reasons that speak against a modular view of the evolution of the mind.

simulated to be understood.[33] The alternative to a modular understanding is that appreciating cognitive adequacy in self or others is performed through metacognitive self-simulation.

Now a major problem for this view is how can a communicator *learn* the cues that predict cognitive adequacy in communicating nonverbal contents. The response is that metacognitive control is learnt as all forms of control are: forward models are constructed based on features of prior exchanges. Such forward models are independently hypothesized as forming the cognitive basis of actions.[34] It is quite plausible that specialized forward models should underwrite informational adequacy, both individually and in cooperation. Communicating systems would use dynamic cues to make predictions of adequacy, and produce practical, on-line evaluations. Some of these cues might be made publicly accessible through gestures and linguistic markers: they would allow participants to establish a (minimally) common evaluation of conversational adequacy. Thus, feelings of understanding, of confusion, of effortful reasoning, etc. can be felt, expressed, or both.

If this analysis is on the right track, engaging in conversation requires metacognitive capacities (exercising simulation), rather than mind-reading capacities, (attributing to someone a mental state). Even though any self-simulation can be *redescribed* in fully conceptual terms, for the purpose of report, justification, etc., it does not need to be. This leads us to claim 2.

2) Metacognitive development is *phylogenetically and ontogenetically* distinct from the development of metarepresentation and mind-reading.

Self-simulation allows to covertly prepare action, evaluate other's as well as own performance. It seems to occur at many different processing levels.[35] It generates primary forms of procedural reflexivity that are later exploited in higher-level, language-based metarepresentational descriptors. Recent findings support the claim that metacognitive self-simulation has a phylogenetic realization earlier than mentalising abilities (reasoning in mental terms about others' intentions and beliefs). Marine mammals as well as monkeys typically fail false belief tasks: they do not seem able to metarepresent (conspecifics' or their own) mental states *as mental states*.[36] On the other hand, they can use endogeneous cues

[33] As stressed in Proust (in print), the needed self-simulation does not have to occur in each occasion. Self-simulation for a task must be a disposition acquired by the recipient, based on prior exercise.
[34] See for example Wolpert et al., 2003.
[35] See Decety et al. (1997). See also Proust (2006).
[36] See Smith et al. (1995), (1997), (1998), (2006).

(analogous to humans' epistemic feelings) to predict/evaluate success and failure in perception and memory. Although developmental research on human children has often supposed that metacognition and metarepresentation develop simultaneously, new research suggests that they might actually be influenced by a third factor, namely conversation.[37] Conversation might both exercise metacognition, by constantly updating memories and stimulating metamemory, a metacognitive capacity; and thereby pave the way for metarepresentation, by offering the semantic vehicles for belief expression. Children might thus learn how to use belief concepts from initially empty labels such as "I believe that - " through what is called 'the ascent routine'.[38] Space does not permit a full reconstruction of how such a development might go. [39] The point is that although social non-human animals are rarely motivated to communicate what they do know, the few species that possess metacognitive capacities are likely to have dedicated somatic markers and epistemic feelings, and therefore might communicate their metacognitive states to others through gestures as humans do.

3) Metacognitive capacities are task-specific rather than domain-specific

We are now in a position to address the question of processing effort that was raised above. Metacognition is *task-specific* because it uses prior responses in a similar task to set a task-related *norm*, and evaluate on its basis any judgment concerning various aspects of observed or anticipated performance on the same task. For example, one knows that one can retrieve a proper name, say, because one has stored facts about one's prior memory retrieval. One knows one's efficiency, temporal course and margin of error in memory retrieval in a practical way, through an epistemic feeling; such procedural knowledge is constantly used in conversation when one has to decide whether it is appropriate to try to remember somebody's proper name.[40] Metarepresentations, on the other hand, do not have this relation to self-evaluation, and are not task-specific. One can report others' beliefs, desires, intentions, as

[37] De Villiers, (2000), Harris et al. (2005).
[38] On the ascent routine, see Evans (1982), Gordon (1996), (Proust (2003).
[39] As I have argued in Proust (in print), a full-fledged, "deep" metarepresentational understanding of "I believe" requires exercising *both* the capacity to metacognize that I believe (with the evaluative/corrective ability associated to it) and the ability to master the concept of belief in a general way, i.e. to apply it in third-person attributions.

[40] See Koriat (2000).

well as sentences and gestures (even outlandish or partly understood) in verbal or gestural terms. These reports are usually said to be "domain-specific" because they are built with mental verbs such as "see", "claim", "believe", "desire" – all concepts that are supposed to be learnt during childhood as part of a theory of mind. One of the differences of the present account thus concerns the scope of metacognitive operations: the kind of input they accept, as well as the states they influence.

On a metacognitive view, processing effort is computed on the basis of stored norms in similar tasks. Conversational tasks however vary substantially from one context to another. There is a kind of effort typical of ordinary conversation, another of a philosophical conference, still another in a court of justice. Given how tired one feels, one can be ready for one and shun the others. But the question is: how can we appreciate this, and use it to select producer's meaning of (1)?

As we know from action theory, it is one thing to launch a command, and another to monitor it.[41] Effort has long been considered to be related to on-line monitoring. According to this monitoring view, the intensity/difficulty of processing is appreciated on the basis of the feedback that it generates. In the light of this "monitoring" view, we should thus complete the ordered sequence/effort theory as follows: the producer and the recipient implicitly agree on the fact that a processing sequence involving few steps counts as "relevant" because it *generates* a feeling of ease of processing.

As Koriat et al. (2006) have shown, however, the feeling of effort might be a joint effect of *control* cues and of *observed feedback*. On this more complex theory, control itself may generate a sense of effort. Merely producing a command, in the speaker, (to start producing a message) might already program the level of effort required to process it. The producer would therefore implicitly know from the command that was set, how complex or deep the sequences are to be, to achieve the required processing. The whole communicative act might thus be influenced right from the start by devoting part of the embodied message to this "effort condition". A significant part of embodied conversational metacognition (through intonation, facial expressions, posture change and various gestures for recruiting more or less attention) seems indeed to have the function of maintaining between speaker and hearer a *similar* allocation of resources to complete the relevant computations. Although we cannot

[41] Action theorists have been the first to examine how a subject might represent "effort" in performing a given action. They have shown that to represent effort you need to associate to a given command its observed effects, which become overtime internal feedback to predict future effort. Efforts performed in representing or thinking can be analysed similarly. A mental task is effortful as compared with other tasks of the same kind. The kind of control that you initially put in a task, as much as the feedback that you receive once commands are sent, jointly determine where you currently are in terms of subjectively felt effort.

develop this theory here, it is clear how such a view deeply affects the very concept of sense of effort: for if effort is predicted right at the control level and can be modulated at will, the producer can regulate the level of effort intensity required for the recipient to grasp what (s)he means (increasing it or decreasing it, as the case requires). If this analysis is correct, conversational metacognition has a fundamental role in establishing the effort involved in achieving relevance.

Section 4 – Conversational metacognition, cooperation and defection

We have proceeded until now under the basic assumption that communication is a form of unrestricted co-operation: we share with others our knowledge about the world, bringing the imagistic force of gestures to complement verbal utterances. We share, in addition, our sense of uncertainty about what is communicated. We express our self-knowledge through conversational metacognition, and we reveal through it our critical hindsight concerning others' communicative productions as well. The basic assumption however cannot be right. Speech being performed *for the sake of* the audience contradicts what is known in nature on the function of communication, which is to serve the producer.[42] Does language constitute an exception to Darwinian selection, by favoring the *recipient* of the information rather than the communicating agent? Evidence suggests that the recipient is not universally offered trustworthy information. Humans as well as non-human animals are selective in their information behavior, and may cheat or retain information when no kin is involved, when reciprocity is not possible or when no status is likely to be gained.[43] Another difficulty for the basic assumption is that embodied speech seems to involve little cost, whereas, in nature, honest signalling is always costly to produce, which is deemed to proceed from an evolutionary pressure on informational manipulation.[44] All these difficulties seem to culminate with the very notion of a gestural-conversational metacognition. Why would someone *want to* make publicly available highly sensitive data, such as one's current self-doubts and evaluations of (in)competence ? Why would one intend to share one's uncertainty about one's knowledge states, and thus become predictable, and thereby manipulable, by others ?

[42] For an exhaustive review of the arguments, see Dessalles (1998).
[43] Palmer (1991), Barrett et al. (2002).
[44] Zahavi & Zahavi (1997).

This difficulty has to do with the fact that conversational metacognition seems by definition to be cooperative, and to be more or less reducible to processes implementing Grice's maxims. Applying Grice's classical analysis to conversational metacognition, we end up with the following story: the intention of the speaker/gesturer is to make manifest his/her own metacognitive comments through the present speech/gesture sequence by having the recipient grasp this metacognitive comment as a result of his/her recognition of the producer's intention to get him/her do so by producing this gesture. We saw above, however, that an analysis based on third-degree intention is too demanding. But it is so not only because it makes human communication a very sophisticated affair; but also because no rational agent would wish to expose his metacognitive states to others, and be constrained by cooperative principles when evaluating what to do next. It is obviously more advantageous in certain cases to pretend, for example, to understand what was expressed, and play-act accordingly (by nodding, etc.) than publicly recognize one's failure as a recipient of the communicative sequence. Section 3 above partly addresses the difficulty, by showing that conversational metacognition does not amount to representing one's mental states; it rather expresses uncertainty about informational adequacy of the current exchange, and constructs a common norm for the effort to be invested in an exchange.

Even in this considerably simplified theory, the problem of self-exposure is still arising: *why would one want to inform another person on one's epistemic adequacy for a given turn?* Can metacognitive transparency be a norm for conversation?

Two important considerations bear on this question. The first brings us back to the biological foundations for human communication. There are several views on the actual function of conversation (transmitting knowledge to kin, planning collective action, making people predictable to each other, publicly denouncing cheaters, ensuring social control, gaining status). On each view, deception can turn communication into exploitation and control. If conversation is primarily in the interest of the producer, (for example, because expressing relevant utterances increases status[45]), the latter should prove the recipient that (s)he deserves his/her trust. If conversation is primarily cooperative, and recipient-oriented, the recipient should be able to indicate whether (and to which degree) his/her informational needs are met by a specific utterance. In both cases, communication should contain preset defenses against abuse: pseudo-informers (liars or misinformed speakers) as well as pseudo-

[45] Dessalles, (1998).

receivers (who pretend to, but actually do not watch or hear) must be detectable in principle.[46] Parasites should also be detected: those that give little and receive much. Reciprocally the overly generous informer should have the capacity to realize that the addressee can make a selfish use of the information conveyed.[47]

The second consideration is that, even if it is conceded that conversation involves divergent interests, and therefore involves forms of competition as modeled by game-theory, it also needs to include *some* amount of cooperation: as we saw above, if basic constraints fail to be fulfilled, communication will not occur.

Metacognitive states or dispositions reflect the instability of any communicative norm between these two boundaries. Metacognition can be misrepresented to others just as first-order contents can be. Therefore conversational metacognition does not need to reflect moment by moment the individual underlying metacognitive feelings and judgments of the participants. But there is a limit to the divergence between individual metacognitive evaluation and its public expression. Beyond that limit, the very possibility of communication evaporates. Even highly competitive participants must extract conversational meaning, by sharing a metacognitive norm of relevance. Other areas of metacognition, however, encompass more troubled waters.

It is interesting here to compare the role of metacognition and of folk logic as defenses against deception. Sperber & Wilson (2002) have suggested that folk logic evolved as such a defense; trustworthy speakers are able to display the logical structure of the arguments leading to a given conclusion. Conversely, cheaters are detected by their inability to pass the test. On this view, folk logic is primarily serving communicational needs. Rhetorics however evolved in turn to convince less agile thinkers on dubious grounds, which in turn (now?) creates selective pressures for finer conceptual expertise.

A similar evolution may apply to metacognition, with the difference that individual metacognition does not seem to be a uniquely social capacity. I have argued elsewhere[48] that metacognition is a regulation directly prompted by increased flexibility in behaviour. Multi-valued regulation indeed creates selective pressures on how to know what one knows and can quickly remember. Relying on a set of covert simulatory processes, individual metacognition

---

[46] Communication with conspecifics is modulated by a tension between trustworthiness and manipulation, as predicted by game theory. See Sober (1994), Hauser (1997), and Proust (2003).

[47] The problem of status theory is that information does not bear its producer on its sleeve. Then a recipient can always use a piece of information without quoting its source and thereby acquire status for himself. This open possibility of stealing status should limit conversation to large groups in ritualized contexts to maintain authorship recognition.

[48] Proust (2006)$_b$

allows to make decisions covertly on the basis of contextual, practical and task-specific self-evaluation To remain viable**,** each organism must work at maintaining the informational quality of its own environment, both internal and external, while selectively restricting the other organisms' access to it.

Now conversational metacognition is not used in deciding how to act (as generally does metacognition), but in communication. Its function is closer to folk logic's : it is to prove to others the value of one's contribution to conversation, the degree of one's conviction or of one's commitment. Such proof is not offered through arguments, but through somatic gestures supposed to display genuine epistemic feelings.

Let us observe that these metacognitive gestures have a potentially high cost (as predicted by honest signalling theories). A fully trustworthy communicator may have to admit failure or incapacity if conversation happens to expose them. In most cases, however, communicators agree to play down the importance of memory lapses and other infelicities.

If this analysis is correct, one could predict that the divergence between individual metacognition and expressed, conversational metacognition will be sensitive to context. Let us imagine the following study. Take a population of researchers, and observe how they make one and the same Powerpoint presentation of their latest work in two types of contexts. In Context 1, they present their work to their collaborators and students. In Context 2, they present it to their competitors at a professional meeting. Let us bet that the two presentations will differ for the quantity of metacognitive gestures expressing self-doubt. How about self-confidence gestures ? I leave it to the readers to come up with their own hypothesis.

Conclusion

The aim of this chapter was primarily methodological and conceptual. It was to show that there is a class of gestures that have a specific metacognitive function, and deserve to be studied as such. We first explored the common properties of metacognitive gestures, and contrasted them with other forms of gestures as well as with other forms of metacognitive engagements. We discussed the issue of the respective functions of cognitive and metacognitive conversational gestures and found interesting parallels and differences, concerning the kind of uncertainty that each kind aims to appreciate and reduce.

Then we examined the alternative case for a first-order, cognitive (rather than metacognitive) approach, claiming that these gestures depend for their acquisition and use on ideomotor or resonance mechanisms rather than on specialized procedures of a different kind.

Although shared emotions might indeed help understand metacognitive gestures, we have shown that they don't suffice to provide a basis for learning how to use them. Metacognitive gestures, it was claimed, presuppose mechanisms of self-simulation, which cannot be acquired by merely simulating another agent. The producer must be able to compare his/her present evaluation of the on-going conversation with a stored norm, accessible through self-simulation and feeling.

We then addressed another popular view, according to which conversational control largely relies on theory of mind and mental reasoning. This view, however, is incompatible with the aptitude of children to converse before they master a theory of mind. We examined the alternative possibility developed by Sperber and Wilson (2002), that relevance might be understood on the basis of a common feeling of effort constraining inferences both at the production and at the reception levels. This interesting but relatively elusive suggestion needs to be explored, and might indeed be subjected to experimental research, as part of a metacognitive experimental apparatus. It is an intriguing possibility that a whole set of metacognitive gestures have the function of calibrating inferential effort among communicators. We ended our discussion with an examination of the evolutionary pressures that are exerted on conversation; how indeed does Machiavellian pressures affect conversational metacognition : how can one ever want to publicly express one's underlying evaluations of one's utterances ? Why should one do it? The response is that doing so is a precondition for communication to be successful in a given range of situations where cooperation is needed. Where extensive cooperation is not required, metacognitive conversational gestures might be used to protect oneself from others' critical evaluations rather than to express one's own.

At this point, no empirical evidence has obviously been collected – whether on conversational metacognitive gestures or on the embodiement for a shared sense of effort. The very concept of conversational metacognition, understood as a set of procedures meant to monitor and control informational adequacy in embodied communication, is entirely new and cries for carefully controlled experiments. It is to be hoped that the present chapter will constitute an invitation for studying it; it would be particularly fruitful to learn how metacognitive gestures develop in children, how deeply they contribute to mutual understanding in adult speakers, and whether and how they are selectively impaired in certain mental pathologies.

APPENDIX : Metacognitive gestures: From function to taxonomy

To explain the existence of metacognitive gestures and their role among other speech gestures, it is important to take a step back, and examine embodied communication as the coupling of two or more dynamic systems.[49] In a dynamic and distributed view of conversation, the kind of control that helps regulate it depends roughly on three sets of constraints.

1) The first offers a general dynamic frame in which exchanges can be performed in a *stable way* in spite of changing conditions concerning content, audience, etc. For example, turn-taking, publicly marked ground sharing, rhythmic embodied attentional patterns, are dynamic organizational principles without which no conversation could occur.

2) The second set determines how an occurrent, or token of, conversation is or remains viable: Gricean maxims, and particularly, the maxim of "relation" – articulated in relevance theory[50] -, state in which conditions gesture and talk can be used successfully to promote *minimizing effort* in communicating one's intentions and recognizing others' intentions. Just imagine what can make a conversation impossible: uttering inferentially unconnected or incomplete sentences, gesturing in a random way, without ever focusing on an audience, or ignoring the audience's informational needs, etc. It is not fully clear yet how Gricean or relevance maxims are operating to ensure cooperation, but some set of mechanisms must ensure that conversation follows a minimally cooperative pattern.

3) The third set of constraints determines the limits in which a system needs to stay to spend its own resources fruitfully. Just as the second set determines the viability conditions of a token of a conversation between two or more participants, the third set determines, at the level of the individual participant, the most viable, ie. the least effortful strategy needed to complete the task.

Actually, this third set of individual constraints might be seen as being at the very basis of the preceding set of collaborative ones, because *the principle of the least collaborative effort* depends asymmetrically on *the principle of the least individual effort*. This last principle, may be applied in two fundamental ways. Either by implicitly learning how to perform the task (when it is recurring in a similar context) or through metacognitive learning

[49] See the contribution of J.S. Jordan, this volume.
[50] See Sperber & Wilson, (1995)

(when the agent has to evaluate the effort needed and his/her occurrent mental dispositions). In cases like this, metacognitive norms (built themselves over time from prior success/failure ratios) instruct agents how to probe their "reflexive uncertainty" (uncertainty about own ability) in various dimensions and how to respond to it (how to decline responding when uncertainty reaches a certain threshold, how to make safe bets, etc.)

Communicational gestures are clearly shaped by the tight interplay of the three sets of constraints. 1) Gestures enhance processing in recipients if they conform to the systems' dynamic patterns; 2) they enrich the communicated contents with non-conceptual representations, with the constraint that this enrichment must fall under cooperation maxims to be at all usable; and finally, 3) gestures must respond to metacognitive worries: they should allow degrees of belief uncertainty and of commitment to be conveyed; they should help predict the dynamics of knowledge acquisition between the participants; they should provide comments on the *quality* of shared information and the resulting acceptability of new proposals.

Is it fruitful, on the basis of these considerations, to set ourselves the task of providing a list of the various metacognitive gestures (associating a gesture with a functional role)? Such a project would not only require collecting videotaped evidence in various illocutionary situations and cultures, which is at present not done on any significant scale.[51] It would presuppose, more radically, that such a principled taxonomic organization exists. One might think that speech act theory offers a taxonomy of utterances, on which a taxonomy of metacognitive gestures could be based. Granting that each type of felicity conditions can be violated, metacognitive gestures might then be categorized as a sensitive evaluation of a particular felicity condition for a speech act. For example, various uses of pointing would be associated with various justifications (or infractions) concerning reference. Requests for information should prompt gestures representing various degrees of anticipated helplessness, confusion or ignorance, etc.

It can first be objected to this project, however, that the standard felicity conditions do not exhaust the range of evaluative dimensions along which metacognitive gestures may be classified (for example, social, moral and political norms might affect gesture production and comprehension). Second, it is generally accepted that conversational gestures cross illocutionary boundaries as much as words do: there is little hope to see dedicated illocutionary metacognitive gestures. Gesture meanings are more often inferred than coded,

[51] Eibl-Eibesfeldt, I., 1974. Similarities and differences between cultures in expressive movements. In Weitz, S. (ed.) , *Non-Verbal Communication*. Oxford, Oxford University Press.

and, if coded, are produced in a complex interplay with inference, as is clearly the case for pointing. The very project of a taxonomy, understood as a clear association between gesture and function, seems hopeless.

Aside from any claim to taxonomy, an interesting question that received relatively little attention until now,[52] is whether metacognitive gestures are more often found with the role of marking the degree of illocutionary force in a given speech act. Assertives should involve ways of expressing one's subjective degree of belief. Requests for information should prompt gestures representing various degrees of anticipated helplessness, confusion or ignorance (one can predict that other kinds of requests should involve much less metacognitive comments). Promises might involve subtle gestural-postural indications on the degree of commitment.[53] Declaratives and expressives might involve gestures displaying self-awareness of performing them with a wide array of possibly contradictory feelings and self-doubt.[54] (In Section 4 above, we saw how these displays pose an interesting, but solvable puzzle to a view of communication where cooperation should not develop at the detriment of individual interests)

Finally a gesture taxonomy cannot be built on a purely individualistic basis. As we noted earlier, metacognitive gestures involve more than an individual sense of ability as made manifest by a participant. Accordingly, conversational analysts often emphasize that utterances and gestures make sense not as single units, but as dynamical entities involving adjacency pairs (Schegloff, 1988). An adjacency pair is a sequence that contains two utterances produced in succession by different speakers. You don't express your epistemic state independently of the person your are talking to and of the task at hand. The two functions of metacognitive gestures examined in section 1 have to be spelled out in this interactive, highly contextual, framework. Metacognitive gestures are meant to be grasped by a recipient (in the "recipient-oriented" function), or they frame the strategy of communicating contents to someone in particular (in the "speaker-oriented" function). In embodied conversational metacognition, participant A may express uncertainty relative to his/her ability to make a true assertion through a gesture or a facial expression (alternatively: to express the degree of his/her commitment to follow a promise, or the depth of his/her regret for a past cognitive failure etc). But whether (s)he does it, and does it with gestures and facial

[52] With some notable exceptions : for example, Kendon (1995) shows that Napolitan conversational gestures, such as mano a borsa (« continued disagreement with the other speaker »), mani giunte (« the premise is an undeniable fact »), or ring (« this piece of information is correct », express complex speech acts which also have an important metacognitive component. See also Poggi (2002) and Poggi & Pelachaud (1998).

[53] Self-grooming, fidgetting, might be strategies suggesting less than whole-hearted commitment.

[54] Facial expression allows to present « mixed feelings » : one can, e.g. express regret in a triumphant way. Again here, we are concerned with intentional expressions of emotion, not with natural signs associated with representations.

expressions of *this* degree and with *this* emotion, depends on the social context and on the recipient's attitude. Participant B will produce in turn an embodied response in which he either accepts or rejects the metacognitive comment displayed by A's gesture. For example, if A produces an assertion displaying the feeling of currently mastering inadequately some content (through an intentional hesitation in speech, "helplessness" gestures or a specific intonation pattern), B may either *accept* A's expressed feeling of not knowing (typically by frowning) or *reject it* by restoring A's attributed "competent knower" status and encourage A to say more (typically by a gesture of both hands extracting something from A). The important aspect in studying "metacognitive pairs" such as these is to examine how they are elicited in different contexts, some facilitating metacognitive avowals, some on the contrary inviting their suppression or their misrepresentation.

To summarize, when they are analysed in the contextual, dynamic conditions of their production, metacognitive gestures should have a varied expression, reflecting the type and degree of cooperation involved, the social roles of the participants, and the importance of the collective goal(s). Far from being the expression of unchanging levels of self-knowledge for a given illocutionary situation, metacognitive gestural communication should be affected by the recipient(s), and be sensitive to interaction and past history of the communicating pair (or group).

References

Alibali, M.W., Kita, S. & Young, A. 2000. Gesture and the process of speech production: we think, therefore we gesture. *Language and Cognitive Processes*, 15 (), 593-613.

Arbib, M.A. (ed.). 2006. *Action to language via the Mirror Neuron System*, Cambridge, Cambridge University Press.

Austin, J.L/1962. *How to do things with words*, Cambridge, Mass.: Harvard University Press.

Baron-Cohen, S. 1995.*Mindblindness ;An Essay on Autism and Theory of Mind*, Cambridge, MIT Press, Bradford Book.

Barrett, L, Dunban, R. & Lycett, J. 2002. *Human evolutionary psychology*. Houndmills, UK : Palgrave Macmillan.

Barsalou, L. 1999. Perceptual symbol systems. *Behavioral and Brain Sciences,* 22:4, 577–660.

Bavelas, J. B., & Chovil, N. 2000. Visible acts of meaning. An integrated message model of language use in face-to-face dialogue. *Journal of Language and Social Psychology,* 19, 163-194

Bavelas, J. B., Chovil, N., Coates, L., and Roe, L. 1995. Gestures specialized for dialogue. *Personality and Social Psychology Bulletin,* 21, 394-405.

Bavelas, J.B., & Chovil, N. 2006. Hand gestures and facial displays as part of language use in face-to-face dialogue. In V. Manusov & M. Patterson (Eds*.), Handbook of nonverbal communication* (pp. 97-115). Thousand Oaks, CA: Sage.

Bavelas, J. B., & Gerwing, J. 2007. Conversational hand gestures and facial displays in face-to-face dialogue. K. Fiedler (Ed.), *Social communication* (pp. 283-308). New York: Psychology Press.

Carpenter, M., Nagell, J., Tomasello,M. Butterworth, G. & Moore, Ch. 1998. Social Cognition, Joint Attention, and Communicative Competence from 9 to 15 Months of Age*, Monographs of the Society for Research in Child Development*, Vol. 63, 4, 1-174.

Clark, H.H. & Foxtree, J.E. 2002.Using *uh* and *um* in spontaneous speaking. *Cognition*, 84; 73-111.

Clark, H.H. &. Krych, M.A. 2004. Speaking while monitoring addressees for understanding Language, *Journal of Memory and Language,* 50 (2004) 62–81.

Clark, H.H. & Wilkes-Gibbs, D. 1986. Referring as a collaborative process. *Cognition*, 22, 1-39.

Crowder, E.M. 1996. Gestures at work in sense-making science talk. *The Journal of the Learning Sciences*. 5, 3, 173-208.

De Villiers, J. 2000. Language and theory of mind : what are the developmental relationships? In S. Baron-Cohen, *Understanding other minds*, Second Edition, Oxford, Oxford University Press,83-123.

Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F. & Fazio, F.: 1997, Brain activity during observation of action. Influence of action content and subject's strategy, *Brain,* 120, 1763-1777.

Dessalles, J.L. 1998. Altruism, Status, and the Origin of Relevance, In J. R. Hurford, M. Studdert-Kennedy & C. Knight (eds), *Approaches to the Evolution of Language: Social and Cognitive Bases*. Cambridge University Press.

Duncan, S. 2006. *McNeill coding manual*. Chicago: University of Chicago

Eibl-Eibesfeldt, I., 1974. Similarities and differences between cultures in expressive movements. In Weitz, S. (ed.) , *Non-Verbal Communication*. Oxford, Oxford University Press.

Ekman,P. & Friesen, W. V. 1972. Hand Movements, *Journal of Communication*, 22, 4, 353-374.

Ekman,P. 1979. About brows : Emotional and conversational signals. In J. Aschoof, M. von Cranach K. Foppa, W. Lepenies & D. Ploog (eds*.), Human ethology : Claims and limits of a new discipline*, 169-248, Cambridge, Cambridge University Press.

Evans, G.: 1982, *The Varieties of Reference, Oxford*, Clarendon Press.

Franco, F. 2005. Infant pointing. In N. Eilan Ch. Hoerrl T. McCormack & J. Roessler*, Joint attention: communication and other minds,* Oxford, Oxford University Press, 129-164.

Gerwing, J. & Bavelas, J. 2004. Linguistic influences on gesture's form, *Gesture*, 4, 2, 157-195.

Goldin-Meadow, S. 2003. *Hearing gesture. How our hands help us think*. Cambridge: Harvard University Press.

Goldin-Meadow, S., Alibali, M.W. & Church, R.B. 1993. Transitions in concept acquisition:using the hand to read the mind. *Psychological Review*, 100, 279-297.

Goldin-Meadow, S. & McNeill, D. 1999. The role of gesture and mimetic representation in making language the province of speech. In M. Corballis & St. Lea (eds.) *The Descent of Mind*, Oxford: Oxford University Press, 155-172.

Gordon, R.M.: 1996, 'Radical' simulationism, in Carruthers, P., & Smith, P.K., (eds.), *Theories of Theories of Mind,* Cambridge, Cambridge University Press, 11-21.

Grice, P. 1989. *Studies in the way of words*. Cambridge, Harvard University Press.

Hadar, U. 1989. Two types of gestures and their role in speech production. *Journal of Language and Social Psychology*. 8, 221-228.

Hadwin, J., Baron-Cohen, S., Howlin, P. & Hill, K. 1997. Does Teaching Theory of Mind Have an Effect on the Ability to Develop Conversation in Children with Autism? *Journal of Autism and Developmental Disorders* 27, 5, 519-537

Harris, P.L., Rosnay, M. (de), Pons, F. 2005. Language and Children's Understanding of Mental States. *Current Directions in Psychological Science* 14 (2), 69–73.

Jacob, P. & Jeannerod, M. 2005. The motor theory of social cognition : a critique. *Trends in Cognitive Sciences,* 9, 1, 21-25.

Jacobs, N. & Garnham, A. 2007. The role of conversational hand gestures in a narrative task. *Journal of Memory and Language*. 56. 291-303.

Jeannerod, M. 1999. The 25th Bartlett lecture.To act or not to act: perspectives on the representation of actions. *The Quarterly Journal of Experimental Psychology, Human Experimental Psychology,* 52A, 1-29.

Kendon, A. 1995. Gestures as illocutionary and discourse structure markers in Southern Italian Conversation. *Journal of Pragmatics*, 23, 247-279.

Koriat, A. 2000, The Feeling of Knowing: some metatheoretical Implications for Consciousness and Control. *Consciousness and Cognition*, 9, 149-171.

Koriat, A. Ma'ayan & H. Nussinson, R. 2006. The Intricate Relationships Between Monitoring and Control in Metacognition: Lessons for the Cause-and effect relation between subjective experience and behavior, *Journal of Experimental Psychology: General*, 135, 1, 36 – 69.

Krauss, R.M. 1998. Why do we gesture when we speak *? Current Directions of Psychological Science*, 7, 54-60.

Levelt, W. J. 1983. Monitoring and self-repair in speech. *Cognition*. 14(1):41-104.

McNeill, D. & Duncan, S. 2000. Growth points in thinking-for-speaking, in McNeill (ed.), *Language and Gesture*, Cambridge, Cambridge University Press : 141-161.

McNeill, D. 1992. *Hand and Mind: What gestures reveal about thought*. Chicago, Chicago University Press.

McNeill, D. 2005. *Gesture & Thought*. Chicago, Chicago University Press.

Melinger, A. & Levelt, W.J.M. 2004. Gesture and the communicative intention of the speaker. *Gesture*, 4 :2, 119-141.

Millikan, R. 1993. *White Queen Psychology and Other Essays for Alice.* Cambridge: Bradford Books.

Özyürek, A. 2002. Do speakers design their cospeech gestures for their addressees ? The effects of addressee location on representational gestures. *Journal of Memory and Language,* 46, 688-704.

Palmer, C.T.1991. Kin selection, reciprocal altruism and information sharing among marine lobstermen, *Ethology and Sociobiology*, 12: 221-235.

Poggi I. 2002.Symbolic gestures: The case of the Italian gestionary, *Gesture*, Volume 2, Number 1, 2002, pp. 71-98(28)

Poggi, I. & Pelachaud, C. 2002. Performative faces, *Speech Communication*, 26,5-21.

Prinz, W. 1997. Perception and action planning. *European Journal of Cognitive Psychology*, 9 ; 129-154.

Proust, J. 2003. Can radical theories of simulation explain mental concept acquisition ? in J.

Dokic & J. Proust (eds), *Simulation and knowledge of action,* Amsterdam : John Benjamins, 201-228.

Proust, J.: 2006_a, Agency in schizophrenics from a control theory viewpoint, in W. Prinz & N. Sebanz (eds.), *Disorders of volition*, Cambridge, MIT Press, 2006, 87-118.

Proust, J.: 2006_b, Rationality and metacognition in non-human animals, in S. Hurley & M. Nudds (eds.), *Rational Animals* ?, Oxford, Oxford University Press, 2006.

Proust, J.: in $print_a$, What is a mental function ? In A. Brenner & J. Gayon (eds.), *French Philosophy of Science*, *Boston Studies in the Philosophy of Science*.

Proust, J.: in $print_b$, Is there a sense of agency for thought? In L. O'Brien ed. , *Mental action*, Oxford, Oxford University Press.

Proust, J. in $print_c$. Metacognition and metarepresentation:Is a self-directed theory of mind a precondition for metacognition ? *Synthese*

Reder, L.M. (ed.) 1996. *Implicit memory and metacognition*, Hillsdale : Lawrence Erlbaum Associates.

Rizzolatti, G. & Arbib, M.A., 1998. Language within our grasp. *Trends in Neuroscience*, 21, 188-194.

Samuels, R. 1998. Evolutionary Psychology and the Massive Modularity Hypothesis, *The British Journal for the Philosophy of Science*, 49, 4, 575-602.

Schegloff, E.A. 1984. On some gestures' relation to talk, in J. M. Atkinson & J. Heritage, *Structures of Social Action, Studies in Conversation Analysis*, Cambridge, Cambridge University Press, 266-296.

Schegloff, E.A. 1988. Description in the Social Sciences I: Talk-in-Interaction." *Papers in Pragmatics* 2,1-24.

Smith, J. D., Schull, J., Strote, J., McGee, K., Egnor, R. & Erb, L.: 1995, The uncertain response in the bottlenosed dolphin *Tursiops truncatus Journal of Experimental Psychology: General,* 124, 391-408.

Smith, J. D., Shields, W. E., Allendoerfer, K. R., and Washburn, D. A. : 1998, 'Memory monitoring by animals and humans'. *Journal of Experimental Psychology: General,* 127: 227-250.

Smith, J. D., Shields, W. E., Schull, J., and Washburn, D. A.: 1997, 'The uncertain response in humans and animals'. *Cognition*, 62: 75-97.

Smith, J.D., Beran, M.J., Redford, J.S., Washburn, D.A. 2006. Dissociating Uncertainty Responses and Reinforcement Signals in the Comparative Study of Uncertainty Monitoring, *Journal of Experimental Psychology*, 135, 2, 282-297.

Sperber, D. 2000. Metarepresentations in an evolutionary perspective, in D. Sperber (ed.), *Metarepresentations: A Multidisciplinary Perspective*. Oxford University Press, pp.117-137.

Sperber, D. & Wilson, D. 1986/1995, *Relevance: Communication and Cognition*. Blackwell

Sperber, D. & Wilson, D. 2002, Pragmatics, Modularity and Mind-reading, *Mind and Language,* 17. 3-23.

Tager-Flusberg, H. 1996. Brief report: Current theory and research on language and communication in autism, *Journal of Autism and Developmental Disorders*, 26, 2, 169-172

Wolpert, D.M., Doya, K., & Kawato, M. 2003. A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London*, B 358, 593-602.

Zahavi, A. & Zahavi, A. 1997. *The handicap principle,* Oxford, Oxford University Press.