『語用論研究』第 13 号(2011 年) pp. 1-19

A Relevance-Theoretic Account of the Evolution of Implicit Communication

Anne Reboul
Laboratory on Language, the Brain and Cognition (L2C2, CNRS FRE3406),
University Claude Bernard-Lyon 1

Implicit communication is the single feature of human linguistic communication absent from any other animal communication system. The recent argumentative theory of reasoning, which links the evolution of reasoning abilities in humankind to the evolution of linguistic communication suggests that linguistic communication evolved to allow humans to manipulate each other, and that communicating things implicitly allowed them to hide their manipulative intentions. Manipulation in that sense is compatible with Gricean cooperation.

Keywords: implicit communication, Gricean cooperation, manipulation, intention, argumentation

1. Introduction

The question of the evolution of language, after having provoked a great deal of interest in the eighteenth and nineteenth centuries, underwent a long decline during the late nineteenth and most of the twentieth century, before being revived by Pinker and Bloom in 1990. This triggered a vast number of papers and books (see Bickerton 1996, Deacon 1997, Dunbar 1996, 2004, Fitch 2010, Hauser 1996, Jackendoff 2003, Mithen 2006, Knight, Studert-Kennedy and Hurford 2000, Tomasello 1999, 2008, 2009, Reboul 2007, to mention only a few books), most of which have quoted Hockett's (1963) classical list of thirteen "essential features of language, which make it unique among animal communication systems:

- Language uses the *vocal auditory channel*, with three consequences which are also listed by Hockett among his thirteen essential features: *broadcast transmission* (the utterance can be perceived not only by the addressee, but by everyone in the vicinity), *rapid fading* and *total feedback* (speakers hear what they say);
- Language has *interchangeability* (the speaker can become the hearer and *vice versa*);
- Language has *specialization* (it implies specific encoding-decoding processes);
- Language has *semanticity* (utterances are meaningful);
- Language has *arbitrariness* (linguistic meaning is not generally based on iconic links between morphemes and their referents);

- Language has *discreteness* (utterances differ in their parts, not as wholes);
- Language has *displacement* (it is possible to speak of absent or even of non-existent objects);
- Language has *productivity* (or *generativity*: there is no limit in principle to the number of sentences, which can be produced in a given language);
- Language has *duality of patterning* (at the phonological level, yielding morphemes, and at the syntactic level, yielding sentences);
- Language is *traditionally transmitted* (regardless of whether or not there is a biologically based language system, specific languages have to be learnt).

It seems, however, that Hockett missed an important feature of language: *implicit communication*.[1]

What is more, it should be noted that, though the whole set of features listed by Hockett is indeed specific to human language and not to be found in any other known animal communication system, this is not the case for each feature taken in isolation. Indeed, all of them seem to be found in one species or another (see Fitch 2010). But the core feature of human linguistic communication that was absent from Hockett's list, i.e., implicit communication, is unique in the sense that, as far as is now known, no animal communication system manifests implicit communication. The very fact that implicit communication is the unique feature that is specific to human linguistic communication should make it central to any account of linguistic evolution in that it does not seem to have evolved from anything else.[2] However, despite its uniqueness, it is never mentioned. It is my aim in the present paper to remedy that omission and to propose some tentative ideas as to why language (and only language) allows implicit communication.

The paper is organized as follows: I begin (in §2) by showing that a Neo-Gricean scenario for the evolution of implicit communication is implausible, given that implicit communication is far from automatic. I then remind the reader of the original Gricean

[1] Implicit communication occurs whenever what is communicated in an utterance is different from what is said (in the standard Gricean sense).

[2] Evolution is here to be understood in the neo-Darwinian way of a feature occurring due either to the fact that it offers an adaptive advantage (the individuals with the feature will leave more offspring than those without it) or through genetic drift (where it is neutral). In the first case, the feature may have been selected for itself or may be a byproduct of another feature which itself is an adaptation (so-called exaptation). My goal here is not to give a complete account of the evolution of implicit communication, but rather to indicate why implicit communication is advantageous, hence can be adaptive, whether it is an exaptation or an adaptation. One reviewer remarked that it is inferential communication in general (i.e., content accessed through inferences) rather than implicit communication (in the sense of note 1) that is specific to language. This is right, but in keeping with the remarks above, I take it that the evolution of implicit communication could precisely be parasitic on the pre-existing human abilities for using inferences in communication.

definition of non-natural meaning, which leads directly to the principle of cooperation for the recovery of speaker's meaning (§3). I introduce the argumentative theory of reasoning (§4) arguing that it presupposes a coevolution between language and reasoning abilities and that it opens the door to the possibility that language evolved for manipulative rather than purely cooperative reasons. In §5, I defend manipulative accounts against some objections, relying on Krebs and Dawkins' (1984) classical paper. I then outline a manipulative scenario for the evolution of implicit communication (§6), before discussing further reasons to link argumentation and manipulation (§7). Finally (§8), I outline a multi-layered account of collaboration, argumentation, manipulation and cooperation.

2. Is a minimax account possible?

Minimax accounts are basically economic principles to the effect that cost should be *minimized*, while benefits should be *maximized*. As Horn (2004) pointed out, minimax accounts, though usually not presented in such terms, are frequent in semantics and pragmatics:

- classical rhetorical recommendations (from Aristotle's *Rhetorics* on) encourage speakers or writers to be as concise and clear as possible;
- Post-Gricean accounts, such as Relevance Theory (Sperber and Wilson 1995), are clearly also instances of minimax; and
- Neo-Gricean accounts, such as Horn's and Levinson's (2000), are other instances.

There is, however, a major difference between Post-Gricean accounts and Neo-Gricean accounts: clearly the Relevance Principle concerns the costs and benefits *for the hearer*, while Neo-Gricean accounts aim at balancing costs and benefits *for both the speaker and the hearer*. I will concentrate here on Horn's account, because he explicitly articulates it in minimax terms. Horn's account concerns *Generalized Conversational Implicatures*.

Grice (1989) distinguished between *conventional* implicatures, which are part of the conventional meaning of the utterance and cannot be cancelled without contradiction, and *conversational* implicatures, which are not part of the conventional meaning of the utterance, and are thus cancellable. However, as Grice himself recognized, some conversational implicatures have a conventional component in the sense that they seem triggered by lexical items, though they are nevertheless cancellable. He called these *Generalized Conversational Implicatures*.

Horn's minimax account (Horn 2004: 13) is based on the Q- and R-principles (which are not among theGricean maxims):

- *Q-principle*: say as much as you can modulo Quality[3] and R;
- *R-principle*: say no more than you must, modulo Q.

The Q-principle concerns the hearer and is a lower-bounding guarantee of sufficiency of informative content, while the R-principle concerns the speaker and is an upper-bounding correlate of the law of Least Effort, enjoining *minimization of form*. Thus, Horn insists on *form* as well as *content* in the generation of implicit meaning. This is especially clear in his discussion of categorical sentences (Horn 2004: 11):[4]

A: All/Every F is G.
E: No F is G.
I: Some F is G.
O: Not every F is G (Some F is not G).

As is well-known, A implies I and O is the denial of A (it contradicts A); as well, I implicates O. On Horn's account, it is because the "basic forms" (i.e., A/E) are both more informative and *briefer* than I/O that I implicates O. I will not discuss Horn's account as such (but see Noveck and Sperber 2007, Reboul 2004), but will rather discuss the possibility of giving a Neo-Gricean minimax account of the emergence of implicit communication along the following lines:

> Satisfying both Q- and R-principles will necessarily lead to the emergence of implicit communication as a result of the pressure for making profitable communication.

In other words, implicit communication exists *as a result of the pressure for conveying as much content as possible from as parsimonious as possible a linguistic form.*

There would be several problems with this proposal, the first and least important of which being that it would only account for a small part of implicit communication, i.e., Generalized Conversational Implicatures. It is indeed hard to see how it could account for presuppositions, or particularized conversational implicatures, though it might be argued that it could account for conventional implicatures (through a grammaticalization process). There is however another and more important problem:[5] Neo-Gricean accounts in general, and Horn's in particular, seem to reduce all the cost on the hearer side to the cost of processing the linguistic form, but it is not clear that this is a correct assessment of what happens in implicit communication. Indeed, experimental studies

---

[3] The Gricean maxim of Quality enjoins sincerity (*say what you believe to be true*).

[4] Categorical sentences are the sentences found in syllogisms and were identified by Aristotle who placed them in his well-known *square of oppositions* (see Aristotle, *On Interpretation* §6–7).

[5] Indeed, this problem concerns not only a putative Neo-Gricean account of the emergence of implicit communication, but also the existent Neo-Gricean accounts of Generalized Conversational Implicatures.

on Generalized Conversational Implicatures[6] have shown that they are indeed quite costly to generate (and the implicatures related to categorical sentences have been extensively studied), and this leads to their non-generation in an important proportion of cases, up to 40% (for a review of this experimental work, see Noveck et Sperber 2007, and for a discussion, see Reboul 2008). Given these data, which have been replicated repeatedly in experimental studies, conveying more content with fewer or shorter words is far from being ensured despite the Q- and R-principles. But if this is the case, it is not clear that minimax principles can successfully account for the emergence and the (continuing) existence of implicit communication.

By relying on principles such as Q- and R-, Neo-Gricean accounts avoid having recourse to intentions to account for implicit communication. Given however that the hope of accounting for implicit communication on the basis of a minimax analysis encounters major difficulties, let us now turn to Gricean and Post-Gricean intentional accounts.

3. Intentional accounts of linguistic communication

In his well-known paper, first published in 1957, "Meaning", Grice (1989: 220) introduced the notion of *non-natural meaning* (or *meaning*$_{NN}$), which he defined as follows:

> "A meant$_{NN}$ something by x" is roughly equivalent to "A intended the utterance of x to produce some effect in an audience by means of the recognition of this intention"

This definition is usually interpreted as implying two intentions, a *primary* intention to produce an effect in an audience, and a *secondary* intention that this effect be produced by the recognition of the primary intention.

Similarly, in Relevance Theory, any (human) act of communication is subtended by two intentions:

> *Informative intention*: "To make manifest or more manifest to the audience a set of assumptions *I*" (Sperber & Wilson 1995: 58);
> *Communicative intention*: "To make mutually manifest to audience and communicator that the communicator has this informative intention" (Sperber & Wilson 1995: 60–61).

Thus Gricean and Post-Gricean accounts are not only intentional accounts, but incorporate two embedded intentional levels. This raises the question of why the second em-

[6] Most of the experimental work on implicit communication has concentrated on scalar implicatures, which are the prototype of Generalized Conversational Implicatures.

bedded intention (the secondary intention in the Gricean account and the communicative intention in the Relevance-Theoretic one) is necessary at all.

The answer to that question is that the second intention is necessary to account for implicit communication (see Reboul & Moeschler 1998). In other words, if linguistic communication were codic, a single intention (the primary or informative intention) would be enough. The second intention is there because it plays a role in the recovery of implicitly communicated content: in the Gricean account, through the recognition of the first intention; in the Relevance-theoretic one, because it is the justification for the Communicative Principle of Relevance and the least effort heuristic (see below).

That the second intentional level is not necessary for codic communication can be seen from animal communication, which, as far as is known, is entirely codic, though, in some cases at least, it is intentional, as shown by audience effects. For instance, alarm calls are not produced automatically when the communicator sees a predator, but depend (i) on the presence of an audience, and (ii) on the fact that the audience is appropriate (which usually means that it includes kin; see Hauser 1996, Cheney and Seyfarth 1990). In such cases, the communicator has the intention of producing an effect in its audience (on a Gricean account) or has an informative intention (on a Relevance-theoretic one). And, emphatically, the effect on the audience *does not* depend on its recognition of the communicator's intention. In other words, success in codic communication *does not* depend on either the communicator or the audience having a mind-reading ability (the ability of attributing intentions to others) (see Reboul 2007).

Second intentions come into play when there is a gap between *speaker meaning* and *sentence meaning*. (This is where Searle's (1969) understanding of meaning$_{NN}$ went astray, when he reformulated the secondary intention and incorporated conventional meaning in it). On the Gricean account, in implicit communication, the second intention (that the effect is produced in the audience by the recognition of the primary intention) is necessary to close the gap between speaker meaning and semantic meaning.[7] In Relevance Theory, the communicative intention is what puts the *ostensive* in "ostensive-inferential communication". The fact that linguistic communication is ostensive is the justification for the Communicative Principle of Relevance:

> "Every ostensive stimulus conveys a presumption of its own optimal relevance" (Sperber & Wilson 1995: 260).

And the Communicative Principle of Relevance triggers the least effort heuristic, which is the basic mechanism in pragmatic inferences in Relevance Theory.

Some criticism has been leveled against double accounts of linguistic communication. In one such endeavor, Glüer and Pagin (2003) have argued that double-intentional

[7] Incidentally, this is where the link between the first Grice (the Grice of meaning$_{NN}$) and the second Grice (the Grice of Logic of conversation) comes in.

accounts of linguistic communication presuppose a sophisticated mindreading ability and that, thus, any cases of individuals who are without such abilities but are nevertheless able to communicate would constitute a counter-argument against such theories. They explicitly targeted Gricean and Post-Gricean accounts, taking as a counter-example those autistic or Asperger patients who have sufficient linguistic abilities to communicate. Basically, Glüer & Pagin's argument is of the kind known as *epiphenomenal arguments* in philosophy.[8] They are based on the comparison of the performances of two populations, one with a given cognitive ability (here mindreading), and one without it, in a behavior that is supposed to require that ability (here linguistic communication). If the two populations have identical performances, then it can be claimed that the cognitive ability is, in fact, *epiphenomal* to the performance of the task. Thus, an epiphenomenal argument can only go through if the two populations are behaviorally indistinguishable in their performance. This, however, is far from being the case for linguistically able autistic and Asperger people. Indeed, their difficulties and peculiarities in linguistic communication (including major difficulties with implicit communication, from indirect speech acts to conversational implicatures, through sarcasm) are part of the continuing major difficulties of autistic and Asperger people in social relationships throughout life (for the complete argument against Glüer and Pagin, see Reboul 2006).

Thus, there seems to be rather strong justification for double-intentional accounts, and, ever since Grice's "Logic of conversation", linguistic communication has been taken to be cooperative, because it is only through cooperation between speaker and hearer (the speaker "tailoring" his utterance to the abilities and knowledge he attributes to the hearer) that the two intentions can be recovered.

4. Reasoning and the evolution of communication

In a series of papers, Sperber and Mercier (see Mercier 2009, Mercier & Sperber in press) have proposed that human reasoning abilities are mainly social, in that they have evolved to allow humans to come to collective decisions by persuading one another that a given course of action is best. This supposes two mechanisms: (a) the production of reasons (on the communicator's side), and (b) the checking of proposed reasons (on the addressee's side). Basically, Mercier and Sperber were interested in explaining why experimental studies have often found human reasoning to be wanting (for a general and accessible review, see Piatelli-Palmarini 1994). Most explanations for these

[8] Epiphenomenal arguments are especially frequent in philosophy of mind. A paradigmatic example is to be found in Chalmers (1996). Chalmers was arguing against functionalist (or reductive) accounts of qualia and was trying to show that qualia do not play a role in the production of behavior. Thus qualia would be *epiphenomal* as far as the production and the interpretation of behavior is concerned.

failures in reasoning that have been proposed are in terms of *dual processes*: the idea is that humans have two highly different systems for problem-solving: one, an *intuitive* system, which is based on *heuristics* (for a good presentation of heuristics, see Gigerenzer 2007), and which is fast, often phylogenetically determined, unconscious and strongly domain-specific; the other, an *analytic*, system that is, by contrast, more in keeping with logical norms of reasoning, and is conscious, linear and not domain-specific. On most dual accounts, errors in reasoning in the experimental studies are explained by the fact that heuristics were used, while they were inappropriate in the context.[9] Some dual account theorists, however, do not charge all errors in reasoning to the heuristic system. For instance, Evans (2007) argues that the reasoning system also plays a part in the often rather poor human reasoning performances, mostly through *satisficing*, that is, being content with a good enough rather than an optimal solution.

Mercier and Sperber do not argue against dual accounts of reasoning. Rather, they concentrate on the so-called analytic system. Based on a detailed analysis of the specific "biases" that have been identified in the reasoning literature, for instance the so-called "egocentric bias" (a preference for one's own beliefs over beliefs communicated by others), they argue that reasoning indeed did not evolve for general problem solving, but that it evolved for *argumentation*. In other words, it is at its best in social contexts in which a collective decision has to be reached, and in which, more often than not, one is trying to persuade others of one's own position. This explains why people are often rather bad at tasks where they are asked to solve a fairly abstract (and usually not very relevant to them) problem on their own. By contrast, performances are often much improved when subjects are asked to come up with a collective solution to a given problem.

I will take for granted that Mercier and Sperber are right in their account of the evolution of reasoning. What is important here is the implication of their account of the evolution of reasoning has for the evolution of public languages. Clearly, if reasoning did evolve for argumentation and persuasion, its evolution was intimately linked to that of public languages. This suggests a co-evolutionary scenario in which the abilities for public languages coevolved with the abilities for reasoning and fulfill the same function, i.e., argumentation. The hypothesis that public language evolved for argumentation is supported by Brandom (2000, 2009) from the radically different position of philosophical inferentialist pragmatism.[10] Brandom claims that one important support for

[9] The main idea, which is not in dispute, is that heuristics usually work fairly well, but that they lead to dysfunctional answers when they are used out of the type of contexts for which they have evolved. Given the usually poor ecological design of experimental studies, it is hardly surprising that heuristics would lead to incorrect answers.

[10] *Pragmatism* should not be confused with *pragmatics*. Pragmatism is an American philosophical trend, which was initiated at the end of the XIXth century under the (rather different) impulses of William James and Charles Sanders Peirce (other well-known names in pragmatism are John Dewey and

this hypothesis is the existence in natural languages of logical connectives. This agrees rather well with at least parts of Mercier and Sperber's proposal, especially if one extends Brandom's observation beyond logical connectives to pragmatic connectives in general (see below §7).

Thus, in summary, the argumentative theory of reasoning is strongly linked to the evolution of language as a public system of communication dedicated to argumentation. It thus opens the possibility that, contrary to general opinion, language evolved not through cooperation, but through manipulation, where manipulation is understood as "actively changing the victim's behaviour" (Krebs & Dawkins 1984, 383).

5. The viability of manipulative accounts of communication

In 1984, Krebs and Dawkins claimed that the only possible basis for the evolution of communication (in any species, including human language) was for the manipulation of others, i.e., to control (or at least influence) their behavior. This does seem to agree with Mercier and Sperber's argumentative theory. However, Mercier and Sperber insist that they want to maintain a cooperative view of language and its evolution, and that manipulative accounts meet with insuperable objections. Additionally, given that the centrality of implicit communication seems to imply a degree of cooperation (see §3), it may seem at first glance that the very existence of implicit communication disproves any manipulative account of linguistic communication. However, I will argue that Krebs and Dawkins are right and that in fact, there is no real contradiction between cooperation and manipulation.

So let us begin with Mercier and Sperber's argument against Krebs and Dawkins' manipulative account: "If the fact of receiving information from someone else is usually damaging for the members of a species, then evolution will quickly make them 'deaf' to this information. (...) Given that individuals who send signals must find some advantage in so doing, this means that communication, to be evolutionarily stable, must be 'honest' in a majority of cases" (Mercier 2009: 16. My translation). This argument, however, does not seem entirely fair to Krebs and Dawkins. As Mercier (2009) indicates there are two versions of the Krebs and Dawkins' paper (the first was published in 1978), and in the second version (1984), the authors noted this problem, and, though Mercier does not discuss it, propose a solution to it. I will not discuss this solution in detail, but I will briefly outline an answer to Mercier and Sperber's criticism. Krebs and Dawkins' proposal rests on a distinction between two kinds of manipulations: manipulations that benefit the manipulator, but harm the "victim"; manipulations, which benefit both the manipulator and the "victim". It should be clear that Mercier's criticism quoted above applies in the first case, but not in the second. Indeed, the second

---

Richard Rorty).

case is perfectly compatible with cooperation.[11] As said above (see §3), any viable account of human communication has to include intentions. Thus, the distinction proposed by Krebs and Dawkins should be reformulated in intentional terms:

- if the communicator intends to produce by his utterance a behavior in his addressee, which (the communicator believes) will benefit himself but will be detrimental to the addressee, then the communicator is engaging in *hostile manipulation*;
- if the communicator intends to produce by the utterance a behavior in his addressee, which (the communicator believes) will benefit himself, and will be either beneficial or neutral to the addressee, then the communicator is engaging in *non-hostile manipulation.*[12]

It should be clear that non-hostile manipulation is perfectly compatible with cooperation. What is more, there is no objection to a manipulation-based evolutionary account of communication, as long as the manipulation involved is non-hostile. Indeed, I will now claim that such an account yields a natural explanation for the emergence and continuing existence of implicit communication.

6. The evolution of implicit communication

I will now defend a general hypothesis regarding the evolution of implicit communication:[13]

> Implicit communication evolved to facilitate manipulation by allowing communicators to hide their (manipulative) intentions.

---

[11] As a matter of fact, Krebs and Dawkins dub the second kind of manipulation "cooperation". This, however, seems unsatisfactory. Notably, the "cooperation" in question has nothing to do with Gricean cooperation: though Krebs and Dawkins talk about "mindreading", the activity as described by them has nothing to do with attributing mental states to others, but merely denotes the prediction of others' behavior (in other words, *behavior reading*). Additionally, Krebs and Dawkins have nothing to say about implicit communication, while the Gricean cooperative account is clearly dedicated to it. And finally, as I hope to show, though manipulation may not detrimental to the "victim", this does not mean that no deception is involved.

[12] If the communicator, in the second case, is systematically mistaken in believing that the behavior of the addressee will be either beneficial or neutral to the addressee, then his own reputation will probably suffer and he will probably be much less successful in his later manipulative communication. This is independent of the hostile or non-hostile intentions of the communicator, just as the sincerity of the (mistaken) speaker who asserts a falsehood means that he follows the Quality maxim (say what you believe to be true), even though what he says is false.

[13] Regarding presuppositions and conversational implicatures: the distinction between these types of implicit communication is a matter of the evolution of languages, rather than of language.

This will become clear from examples. I will briefly present examples of most major categories of implicit communication,[14] i.e., presupposition, and conversational implicatures (both Generalized and Particularized Implicatures).

Let us begin with Particularized Conversational Implicatures:

(1) A: Do you know where Anne lives?
B: Somewhere in Burgundy,[15] I believe.
pci: *B does not know where exactly Anne lives.*

B may know quite well Anne's exact address, but not want A to write to or visit Anne. By his utterance, B can hide his intention not to give this information to A, or later deny having had this intention.

In Generalized Conversational Implicatures, we find something rather similar:

(2) A: Has Peter finished his homework?
B: Well, he has done *some* of the exercises.
gci: *Peter has not done all of his exercises.*

Suppose that A is Peter's mother and that B is Peter's older sister. By her utterance, B implicates that Peter has not done all of his exercises and, hence, that he has not finished his homework and, for instance, should not be allowed to go with her to the pictures. At the same time, B can insist that, in producing that utterance, she did not mean to prevent Peter from going with her to the pictures.

Let us now turn to our third case of implicit communication, presupposition:

(3) A: I have decided to give the job of manager of the local branch to John.
B: That's an excellent choice, especially now that he has *stopped* drinking.
pp: *John drank.*

B can claim that he has no unfriendly intention toward John and that, indeed, he praised A's choice, even though he hopes by his utterance to change A's mind as to the suitability of John for the post.

I would like now to point out a few general points about these types of implicit communication. First of all, in none of these cases is there any lie *stricto sensu*. For instance, in the case of (1), it is true that Anne lives in Burgundy, though not that B does not know where exactly. In the case of (2), it is true both that Peter has done all his exercises and that he has done some of them (because *all* implies *some*). And in the

[14] In each case, the content implicitly communicated is either the implicature or the presupposition. And, though there has been debate as to whether generalized conversational implicatures and presuppositions are implicit, it seems clear that they are perfectly in keeping with the definition of implicit communication given in note 1 above. I leave aside conventional implicatures, which raise different questions.

[15] Burgundy is a region in France.

case of (3), it is true both that John does not drink (now) and that he used to drink. Second, in all of these examples, the manipulation need not be hostile to the addressee. In (1), it may be the case that Anne would be nasty to A were they to be in contact and that B is trying to protect A from the repercussions of such rude behavior. In (2), B wants to go to the picture with her boyfriend and he cannot stand Peter. B wants to spare her mother's illusion that Peter is universally beloved. In (3), B may be sure that if John is indeed given the job, he will prove a terrible liability to A's company, and wants to protect A from that danger. Third, though there are neither lies nor hostile manipulative intentions in these cases, this does not mean that there is no deception. The deception involved differs from lies in that it does not concern the content that is explicitly communicated by the utterance. Rather it concerns one intention of the speaker. However, the intention that the speaker can deny having or being committed to is neither the primary or secondary intention (in the Gricean account), nor the informative or the communicative intention (in the Relevance-theoretic one). Nevertheless, arguably, in all of the cases examined above, the speaker preserves the possibility to deny that he or she has a specific intention, though that intention is a further intention: the intention to manipulate the addressee, to lead the addressee to do what the speaker wants him or her to do. Basically, what this means is that you can have (Gricean) cooperation within a manipulative wider context.

Let me, as a conclusion to this section, turn back to Mercier's objection to a manipulative account of the evolution of communication, the conclusion of which I reproduce here: "Communication, to be evolutionarily stable, must be *'honest'* in a majority of cases" (my emphasis). The notion of *honesty* is complicated in two ways. First, it is ambiguous between two readings, only the first of which is relevant to Mercier's argument: honesty may be read in an entirely non-psychological (and non-Gricean) way, as describing the fact that the message has *true* content;[16] honesty may be read in a psychological (and Gricean) way, as describing the fact that the message has a content *which the communicator believes to be true.* A second complication has to do with the *scope* of honesty: does it concern the content explicitly communicated, does it also include the content implicitly communicated, and finally, does it concern intentions, and if so, which intentions does it concern?

The reverse of honesty is *deception* and the complications described above related to the notion of honesty have interesting repercussions regarding deception. First of all, just as there is a non-psychological notion of honesty, there also is a non-psychological notion of deception.[17] There is of course a psychological notion of deception (relevant for human communication) and under that reading, *deception* implies the intention to

[16] I will not go into details, but the argument is parallel to that in note 12 above.

[17] It is known as *functional deception* and has been found in non-human primates (see Byrne and Whiten 1988).

deceive, and, again, the scope of *that* intention makes a difference as to the nature of the deception. If the communicator's intention to deceive concerns the explicitly communicated content of his or her utterance, then he or she *lies*. When it concerns the implicitly communicated content (but not the explicitly communicated one), no lie is involved. This is what happens in (1) where B is certainly not lying (the content he explicitly communicated is true), but the conversational implicature is false (B does know Anne's exact address). When the intention to deceive concerns the speaker's intentions (in general), it should be noted that the distinction between non-psychological and psychological readings of honesty and deception vanishes: one cannot be mistaken about one's own intentions. Obviously, again, no lie is involved. However, a further complication occurs. Given double-intentional accounts of linguistic communication, which as we have seen are necessary to account for implicit communication, deception concerning primary and secondary intentions (or, equally, concerning informative and communicative intentions) would be self-defeating and indeed appears to be plainly impossible.

Let me quickly outline why this is so: if a speaker intends to lie, he will not have any chance of success if his intention of lying is recognized by his putative victim. In other words, lies depend for their success on (a) the fact that the speaker's intention to lie is not recognized by his victim; and on (b) the fact that his primary and secondary intentions (or, equally, his informative and communicative intentions) are recognized by his victim. This implies that, in lies, the scope of deception covers both the explicitly communicated content and the speaker's intention to lie, and it also strongly implies that the speaker's intention to lie is neither the primary nor the secondary intention[18] (or equally the informative or the communicative intention). The same thing, *mutatis mutandis*, applies when deception targets implicitly communicated content. In other words, deception regarding such further speaker's intentions will systematically occur whenever deception exists concerning communicated content (whether this is implicitly or explicitly communicated).

The further intentions implied by deception regarding communicated content are however different from the manipulative intentions implied in all of cases (1) to (3), as we shall now see.

7. Argumentation and manipulation

Mercier (2009, 24) discusses the following example: A wants B to help him attack

[18] This was the basis for one major criticism against Grice's meaning$_{NN}$, that is, that it cannot accommodate lies. I think that the criticism is misguided: it confuses primary and secondary intentions with further intentions, which also exist when no lie is made (for a defense of this position, see Reboul 2007). Here the distinction between *proximate* and *ultimate* intentions introduced below (see §8) is indirectly relevant. The criticism against Gricean meaning$_{NN}$ confuses proximate and ultimate intentions: only proximate intentions are involved in meaning$_{NN}$.

a third party, but B is not enthusiastic. Mercier compares two different ways in which A can try and convince B that the attack is a good idea:

(4) a. John has insulted your mother.
b. Let's go and attack John.
c. Let's go and attack John because he has insulted your mother.

In his comparison, Mercier is mainly interested in whether B will or will not use reasoning. Basically, in saying (4a–b), A is giving B a reason (4a) to justify the conclusion (4b). This may trigger a reasoning process in B, evaluating whether (4a) really is a good reason for (4b). However, if A says (4c) (rather than (4a–b)), he decidedly leads B to use reasoning through the use of the word *because*. As Mercier notes, using such linguistic tools (e.g., *thus, then, if ... then*, etc.) encourages reasoning of the evaluative kind. In such cases (with different degrees of expliciteness), we have argumentation (and reasoning) without manipulation: A has the intention to convince B to attack John, but this intention is clearly explicit in both (4a–b) and (4c).

However, as Mercier himself recognizes, this is not necessarily the case, notably because of the so-called egocentric bias. As Mercier notes (2009, 116; my translation): "In the case of beliefs, the mechanisms of coherence checking should not only make us aware of conflicts, it should also promote our own beliefs over communicated beliefs when they conflict (we can call this an *egocentric bias*)". As he points out, this affects inferential mechanisms as well: "When someone tells us something, when we accept it and make inferences based on this information, the resulting conclusions are considered as ours" (Mercier 2009, 117; my translation) and will be more easily accepted than if they had been explicitly communicated. In other words, going back to examples (4a–c), it may well be that A would have greater success if he only said (4a) and let B draw his own conclusion, rather than adding (4b) or simply saying (4c). Mercier adds (2009, 118. My translation): "The less important the communicator's role in the formation of the conclusion by the addressee, the more the addressee will accept the conclusion. Maximally using implicit elements is a risky tactic, because in some cases, the addressee will not draw the intended conclusion, but it can be very efficient". Indeed, given the egocentric bias, when the addressee already has reached a conclusion which is not the one the speaker would like him or her to adopt, the speaker's best chance may be to use implicit rather than explicit communication. Additionally, the very fact that the speaker wants the addressee to adopt a different conclusion would lead the addressee, if he or she knew it, to reject that different conclusion. In other words, the speaker has better chances of success, if he deceives his addressee on his intentions to change the addressee's mind on this specific point. Thus, he has better chances of success if he manipulates his addressee by hiding his further intention. Hence, though in principle argumentation and manipulation are not necessarily linked, practically they will often occur simultaneously.

8. A multi-layered account of collaboration, argumentation, manipulation, and cooperation

Mercier and Sperber (Mercier 2009, Mercier and Sperber in press) note that argumentation most naturally happens when there is a general agreement about the goal of a collaborative action, but agreement still has to be reached as to the best way of attaining this goal. As we have seen in the previous section, argumentation is often accompanied by manipulation, though both argumentation and manipulation suppose successful cooperative communication in the Gricean sense.

At this point, let me introduce some terminological clarifications. Going back to Krebs and Dawkins' initial notion of *manipulation*, it merely stated that manipulation consists in the communicator endeavoring to control (or at least influence) the addressee's behavior to his own advantage. This, it should be clear, is wide enough to include argumentation in Mercier and Sperber's sense. As stated above, manipulation can be *hostile* (leading to behavior that is advantageous to the communicator, but detrimental to the addressee) or *non-hostile* (leading to behavior either advantageous to both or advantageous to the communicator, and neutral to the addressee). And, of course, manipulation can only be the basis for the evolution of communication if it is mainly non-hostile: in other words, hostile manipulation can (and does) occur, but non-hostile manipulation has to be the general case. Krebs and Dawkins assume (as do Mercier and Sperber) that non-hostile manipulation is, quite simply, cooperation. Though Krebs and Dawkins' version is certainly not cooperation in the Gricean sense (see note 11), in humans, non-hostile manipulation seems to correspond to *argumentation* in Mercier and Sperber's sense. *Argumentation* will then be characterized as communication targeting *arguments*, i.e., information whose content is seen by the communicator as a *reason* for a specific behavior that the communicator wants the addressee to perform. *Reasoning* is basically used in the production and evaluation of arguments.

The communicator can use four basic strategies:

i. making his argument *entirely explicit* by *explicitly indicating (using logical and pragmatic connectives) the articulation of the reason and the conclusion* he wants the addressee to draw from it (as in example (4c));
ii. making his argument *largely explicit* by *explicitly communicating the reason and the conclusion* he wants his addressee to draw from it (as in example (4a–b));
iii. making his argument *partly explicit* by *explicitly communicating the reason*, and leaving the addressee to draw the conclusion (as in example (4a));
iv. making his argument *implicit* by *implicitly communicating the reason* and leaving the addressee to draw both the reason and the conclusion (as in examples (1), (2) and (3)).

All of these strategies have advantages and disadvantages:

a. when the communicator chooses the *entirely explicit* strategy, he risks the addressee negatively evaluating the articulation of the reason and the conclusion (considering that the communicator's reasoning as articulated is invalid), rejecting the conclusion, or even rejecting the reason;
b. when the communicator chooses the *largely explicit* strategy, he risks the addressee rejecting the conclusion, or even rejecting the reason;
c. when the communicator chooses the *partly explicit* strategy, he risks the addressee not drawing the conclusion, or rejecting the reason, but given the egocentric bias, he runs less risk of the addressee rejecting the conclusion if the addressee draws it, provided the addressee accepts the reason;
d. when the communicator chooses the *implicit* strategy, he risks the addressee not drawing the reason, or drawing the reason but not the conclusion, but, given the egocentric bias, he runs less risk of the addressee rejecting the reason or the conclusion if the addressee actually draws either or both of them.

In any episode of argumentation, one can distinguish between *proximate intentions* (basically, these are the primary and secondary Gricean intentions, or the informative and communicative intentions) and *ultimate intentions* (these concern the behavior that the communicator intends to induce in his addressee). The four strategies above can be distinguished as to how explicit they are regarding the communicator's ultimate intentions:[19] in the entirely and the largely explicit strategies, they are explicitly communicated; in the partly explicit strategy, they are not communicated, but it may be difficult for the communicator to deny having them, given that he has explicitly communicated the reason; finally, in the implicit strategy, the ultimate intention is not only not communicated, it can be denied, because the reason is not explicitly communicated. In other words, in the implicit strategy, deception (by hiding information) is involved concerning the communicator's ultimate intention.

In other words, these are cases in which there is *non-hostile manipulation (and argumentation) with deception as to the communicator's ultimate intentions*: in this case, arguably, there is *manipulation* in the usual vernacular sense. As said before (see §7), this will occur fairly often and it explains why language allows implicit communication. Thus we come to a short story regarding the evolution of language as the only animal communication system allowing implicit communication: humans, as social animals, depend on one another and more often than not engage in collaborative actions; this requires them to reach agreement not only regarding the goal of these collaborative actions, but also regarding the choice of means; such agreement can only be reached by communication of reasons, i.e., with argumentation; the very fact that most human activity is collaborative opens it to the possibility of hostile manipulation; this has led to

[19] As we saw (see §6), proximate intentions cannot be the object of deception.

the development of preventive mechanisms such as the evaluation of arguments as reasons and to the egocentric bias; the egocentric bias leads to a preference for one's own beliefs and will induce a preference for beliefs which one has reached by oneself; this explains why it may be advantageous for the communicator to use implicit communication: it allows him to induce in his addressee beliefs (i.e., reasons and conclusions) which the addressee having reached them by himself will be more prone to accept and to hide his ultimate intentions regarding the conclusion he wants the addressee to reach as to the future course of her action.

9. Conclusion

If Mercier and Sperber are right regarding both the evolution of reasoning and the argumentative nature of public languages, the existence of implicit communication makes perfect sense. Implicit communication is a way for the communicator to hide the (possibly benevolent) ultimate intentions that he has regarding the addressee's behavior. This allows for widespread (and usually benevolent and non-hostile) manipulation.

To sum up, the basic idea is the following: the uniqueness of implicit communication to language gives a clue as to why linguistic communication evolved. Most current scenarios of language evolution are so-called "social" scenarios, though that rather wide term covers a variety of meanings. However, behind this variety lies the common idea that the main motor of language evolution is cooperation, understood as involving *altruism* in the biological sense, that is, behaviors that are costly to the agent and beneficial to the recipient. In my view, the existence of implicit communication should lead to a more layered view, where manipulation plays as much a role as cooperation in the evolution of linguistic communication. In agreement with the recent theory of argumentation proposed by Sperber and Mercier for the evolution of human reasoning abilities, which strongly suggests that language and reasoning abilities coevolved to allow people to come to an agreement, I propose that implicit communication allows the speaker to hide his intentions to persuade the hearer that the course of action the speaker proposes is better than the one the hearer contemplates (i.e., to manipulate the hearer). The speaker can thus lead the hearer to change his mind, by communicating the relevant information implicitly, rather than explicitly. This does not mean however that implicit communication can only be found in competitive contexts, given the possibility described above of, so to speak, benevolent manipulation. This theory accounts for the very existence of implicit communication and it also explains facts such as the following: implicit communication is often used in polite exchanges, e.g., indirect requests, such as *Could you pass the salt, (please)?*, where a direct request might be seen as rude and where implicit communication allows the speaker to implicitly communicate his request without making his intention that the hearer obey his request too clearly a mark of dominance; implicit communication often communicates not only an (implicit) refusal, but also a general reason for the refusal, thus making it innocuous as it is not seen

as directly linked to the present offer and the person making it. In such cases, manipulation need not be deceptive. In other words, an ability, implicit communication, initially evolved for hiding manipulative intentions, is now also occasionally used without this type of deception.

Thus, under the scenario outlined above (see §8), we get the following layered picture: collaboration → collective decision making → argumentation → manipulation → implicit communication.

References

Aristotle. 2008. *On Interpretation.* In Aristotle (ed.) *Complete Works of Aristotle (Part One),* 204–1403. New York: Rajasir (Kindle Edition).

Aristotle. 2009. *Rhetorics.* In Plato and Aristotle (eds.), *The Works of Plato & Aristotle,* 43391–45867. New York: C&C Web Press (Kindle Edition). (Translation: Benjamin Jowett)

Bickerton, D. 1996. *Language and Human Behavior.* London: Routledge.

Brandom, R. 2000. *Articulating Reasons: An Introduction to Inferentialism.* Cambridge, MA: Harvard University Press.

Brandom, R. 2009. *Reason in Philosophy: Animating Ideas.* Cambridge, MA: Harvard University Press.

Byrne, R. W. and A. Whiten. 1988. "Tactical Deception of Familiar Individuals in Baboons." In R. W. Byrne and A. Whiten (eds.) *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans*, 205–210. Oxford/New York: Oxford University Press.

Chalmers, D. J. 1996. *The Conscious Mind: In Search of a Fundamental Theory.* New York/Oxford: OUP.

Cheney, D. L. and R. M Seyfarth. 1990 *How Monkeys See the World. Inside the Mind of Other Species.* Chicago/London: University of Chicago Press.

Deacon, T. W. 1997.*The Symbolic Species: The Co-Evolution of Language and the Brain.* New York: Norton.

Dunbar, R. 1996. *Grooming, Gossip and the Evolution of Language.* London: Faber & Faber.

Dunbar, R. 2004. *The Human Story: A New History of Mankind's Evolution.* London: Faber & Faber.

Evans, J. St. N. T. 2007. *Hypothetical Thinking: Dual Processes in Reasoning and Judgment.* Hove/New York: Psychological Press.

Fitch, W. T. 2010. *The Evolution of Language.* Cambridge/New York: Cambridge University Press.

Gigerenzer, G. 2007. *Gut Feelings: The Intelligence of the Unconscious.* New York/London: Penguin Books.

Glüer, K. and R. Pagin. 2003. "Meaning Theory and Autistic Speakers." *Mind & Language* 18, 23–51.

Grice, H. P. 1989. *Studies in the Way of Words.* Cambridge, MA: Harvard University Press.

Hauser, M. 1996. *The Evolution of Communication.* Cambridge, MA: MIT Press.

Hockett, C. F. 1963. "The Problem of Universals in Human Language." In J. Greenberg (ed.) *Universals of Language,* 1–29. Cambridge, MA: MIT Press.

Horn, L. 2004. "Implicature." In L. Horn and G. Ward (eds.) *The Handbook of Pragmatics*, 3–28.

Oxford: Blackwell.

Jackendoff, R. 2003. *Foundations of Language: Brain, Meaning, Grammar, Evolution.* Oxford/ New York: Oxford University Press.

Knight, C., M. Studdert-Kennedy, and J. R Hurford. (eds.) 2000. *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form.* Cambridge: Cambridge University Press.

Krebs, J. R. and R. Dawkins. 1984. "Animal Signals: Mind-reading and Manipulation." In J. R. Krebs and N. B. Davies (eds.) *Behavioural Ecology: An Evolutionary Approach*, 380–402. Sunderland, MA: Sinauer Associates.

Levinson, S. 2000. *Presumptive Meanings. The Theory of Generalized Conversational Implicature.* Cambridge, MA: MIT Press.

Mercier, H. (2009) *La Théorie Argumentative du Raisonnement.* Ph.D. Thesis in Social Sciences, Mention Cognitive Sciences, E.H.E.S.S, Paris.

Mercier, H. and D. Sperber. 2011. "Why do Humans Reason? Arguments for an Argumentative Theory." *Behavioral and Brain Sciences* 34:2, 57–74.

Mithen, S. 2006. *The Singing Neanderthals: The Origins of Music, Language, Mind, and Body.* Cambridge, MA: Harvard University Press.

Noveck, I. A. and D. Sperber. 2007. "Why and How of Scalars." In N. Burton-Roberts (ed.) *Advances in Pragmatics.* Basingstoke: Palgrave.

Piatelli-Palmarini, M. 1994. *Inevitable Illusions: How Mistakes of Reason Rule Our Minds.* New York: John Wiley & Sons.

Pinker, S. and P. Bloom. 1990. "Natural Language and Natural Selection." *Behavioral and Brain Sciences* 13:4, 707–784.

Reboul, A. and J. Moeschler. 1998. *La Pragmatique aujourd'hui.* Paris: Le Seuil.

Reboul, A. 2004. "Conversational Implication: Nonce or Generalized." In I. Noveck and D. Sperber (eds.) *Towards Experimental Pragmatics*, 322–333. Basingstoke/New York: Palgrave Press.

Reboul, A. 2006. "HOT Theories of Meaning: the Link between Language and Theory of Mind." *Mind & Language* 21, 587–596.

Reboul, A. 2007a. *Langage et cognition humaine.* Grenoble: Presses Universitaires de Grenoble.

Reboul, A. 2007b. "Does the Gricean Distinction between Natural and Non-Natural Meaning Exhaustively Account for all Instances of Communication." *Pragmatics and Cognition* 15:2, 253–276.

Reboul, A. 2008. "Noel Burton-Roberts (ed.) *Pragmatics* (Palgrave Advances in Linguistics). Basingstoke: Palgrave Macmillan, 2007: Pp xi + 267". *Journal of Linguistics* 44, 519–524.

Searle, J. R. 1969. *Speech Acts.* Cambridge/New York: Cambridge University Press.

Sperber, D. 1994. "The Modularity of Thought and the Epidemiology of Representations." In L. A. Hirschfeld and S. A. Gelman (eds.) *Mapping the Mind: Domain Specificity in Cognition and Culture*, 39–67. New York: Cambridge University Press.

Sperber, D. and D. Wilson. 1995. *Relevance: Communication and Cognition.* 2nd edition. Oxford: Basil Blackwell.

Tomasello, M. 1999. *The Cultural Origins of Human Cognition.* Harvard: Harvard University Press.

Tomasello, M. 2008. *Origins of Human Communication.* Cambridge, MA: MIT Press.

Tomasello, M. 2009. *Why We Cooperate.* Cambridge, MA/London: MIT Press.