Does Linguistic Communication Rest on Inference?

François Recanati

▶ To cite this version:

Does Linguistic Communication Rest on Inference?[*]

François Recanati

Abstract

It is often claimed that, because of semantic underdetermination, one can determine the content of an utterance only by appealing to pragmatic considerations concerning what the speaker means, what his intentions are. This supports 'inferentialism' : the view that, in contrast to perceptual content, communicational content is accessed indirectly, via an inference. As against this view, I argue that primary pragmatic processes (the pragmatic processes that are involved in the determination of truth-conditional content) need not involve an inference from premisses concerning what the speaker can possibly intend by his utterance. Indeed, they need not involve any inference at all : communication, I argue, is as direct as perception.

- Address for correspondence : Institut Jean Nicod, 1bis, avenue de Lowendal, 75007 Paris, France.
- Email : recanati@ehess.fr

1. The Standard Picture

On a widely accepted picture — henceforth the 'standard picture' — linguistic communication exploits two very different sorts of competence, one having to do with the interpretation of linguistic forms, the other with the interpretation of human actions. The recovery of the message communicated by the speaker who uttered a sentence is seen as a hybrid process involving both *semantic interpretation* and *pragmatic interpretation*.

Knowing a language, it is said, is like knowing a system of rules or principles — a 'theory' — by means of which one can deductively establish the truth-conditions of (hence the proposition expressed by) any well-formed sentence of the language. Semantic interpretation is the process whereby an interpreter exploits his or her knowledge of a language, say L, to assign to an arbitrary sentence of L its truth-conditions. This involves (i) determining the meaning of the sentence on the basis of the meanings of its parts and the way they are put together, and (ii) assigning values to indexical expressions in context. (Those assignments are themselves determined by linguistic rules, according to the standard picture. The rules in question — e.g. the rule that 'I' refers to the speaker — constitute the context-independent meaning of indexical expressions.)

Pragmatic interpretation is a totally different process. It is not concerned with language per se, but with human action. When someone acts, whether linguistically or otherwise, there is a reason why she does what she does. To provide an interpretation for the action is to find that reason, that is, to ascribe the agent a particular intention in terms of which we can make sense of the action. Pragmatic interpretation thus construed is characterized by the following properties:

• Pragmatic interpretation is possible only if we presuppose that the agent is *rational*. To interpret an action, we have to make hypotheses concerning the agent's beliefs and

desires; hypotheses in virtue of which it can be deemed rational for the agent to behave as she does.

• Pragmatic interpretation is *defeasible*. The best explanation we can offer for an action given the available evidence can always be overriden if enough new evidence is adduced to account for the subject's behaviour.

• There is *no limit* to the amount of contextual information that can in principle affect pragmatic interpretation. Any piece of information can turn out to be relevant and influence the outcome of pragmatic interpretation.

A particular class of human actions is that of communicative actions. That class is defined by the fact that the intention underlying the action is a communicative intention — an intention such that (arguably) its recognition by the addressee is a necessary and perhaps sufficient condition for its fulfilment. To communicate that *p* is therefore to act in such a way that the addressee will explain one's action by ascribing to the agent the intention to communicate that *p*. For communication to succeed, the addressee must not only understand *that* the agent does what he does in order to communicate something to her; she must also understand *what* the agent tries to communicate. To secure that effect the communicator will do something which will evoke in the adressee's mind that which he wants to communicate. To that purpose the communicator may use icons, or indices, or symbols, that is, conventional signs. It is, of course, symbols that are used when the communicator and the addressee share a common language.

At this point semantic interpretation and pragmatic interpretation make contact with each other. A speech act is an action performed by uttering a sentence in some language L. Let us assume that the sentence has a certain semantic interpretation in L: it means that *p*. Since the speaker utters a sentence which means that *p* and manifests the intention to communicate something to the hearer, one reasonable hypothesis is that he intends to communicate that *p*. If that is the best explanation for the action given the available evidence, the hearer will settle for it and (if that was indeed the speaker's

intention) the communicative intention will be fulfilled: the speaker will have succeeded in communicating that *p* to the hearer. In this case the speech act will be assigned a particular content as a result of pragmatic interpretation; and that content will coincide with the content which comes to be assigned to the sentence as a result of semantic interpretation. That is not really a coincidence, of course; for the semantic interpretation of the sentence was part of the evidence used in pragmatically determining the content of the speech act. But there are cases in which the two contents do not coincide: the sentence means that *p,* but that is not what the speaker means — what he manifestly intends to communicate.

On the standard picture, then, there is a basic distinction between what the sentence says and what the speaker means, even when they coincide. *What the sentence says* is determined by semantic interpretation, that is, deductively and without paying any regard to the speaker's beliefs and desires. Of course, one needs to make sure that the speaker utters what he does as a sentence of L; and that may require a good deal of pragmatic interpretation.[1] But once it is determined that the utterance at issue counts as an utterance of a particular sentence of L, semantic interpretation takes over, and the content of that sentence is mechanically determined.— On the other hand *what the speaker means* is determined by pragmatic interpretation. It relies on a general assessment of the speaker's beliefs and desires, given an overall assumption of rationality. As I pointed out any piece of contextual information may turn out to be relevant to establishing the correct interpretation for the speech act.

2. Inferentialism vs. Anti-Inferentialism

There is a good deal of truth in the standard picture. In particular, the distinction between the two types of interpretation and the idea that both play a role in linguistic communication is important. There is also something deeply wrong with that picture, as

---

[1] This antecedent role of pragmatic interpretation, prior to semantic interpretation, shows that the latter is embedded in the former rather than merely providing an input for it.

we shall see : it ignores the phenomenon of semantic underdetermination (section 3). Before criticizing the standard picture, however, I would like to introduce a debate between two positions, or two attitudes, with respect to linguistic communication construed as involving the two types of interpretation. The debate can be stated in terms of the standard picture, and that's what I am about to do; but it can also be stated in terms of the more refined view we will arrive at after the standard picture has been duly criticized. (My aim in this paper is precisely to show that the debate in question survives the refinement of the standard picture made necessary by the phenomenon of semantic underdetermination.)

The two positions in the debate I call 'inferentialism' and 'anti-inferentialism':

• *Inferentialism* starts by observing that pragmatic interpretation is dominant in linguistic communication. Semantic interpretation is subservient to pragmatic interpretation and provides only part of the evidence used in determining the content of the speech act. Ultimately that content has to be inferred from premisses concerning what the speaker can possibly intend by this utterance in this context. Even if the speaker means exactly what he says, that he does so can only be determined by assessing his intentions. Linguistic communication therefore rests on mind-reading: the content of the speech act can only be ascertained by reflecting on the speaker's mental states and attempting to answer questions such as: Why — for what reason — has the speaker uttered those words? What is he up to? It follows that subjects with deficient mind-reading abilities should be equally deficient in their communicative capacities — a point on which Francesca Happé insisted in her oral communication to the Oxford workshop.

• *Anti-inferentialism* stresses the fact that, in normal cases, semantic interpretation by itself gives us the content of the speech act. Normally the speaker means what he says: to interpret an utterance, therefore, one has only to figure out what the sentence says. If the sentence says that p, one concludes that the speaker means (intends to communicate) that p. Sometimes that conclusion may seem unacceptable. Then, *and*

only then, does a process of reflection upon the speaker's intentions occur. We start wondering why the speaker says what he says when, and only when, there is something wrong with the output of the normal interpretive process, which consists in simply assuming (without reflection) that the speaker means what he says. Only when the unreflective, normal process of interpretation yields weird results does an inference process take place whereby we use evidence concerning the speaker's beliefs and intentions to work out what he means on the basis of what he says.

According to (this version of) anti-inferentialism, semantic interpretation automatically delivers an interpretation of the speech act, without any need for representing the speaker's beliefs and intentions. To be sure, the output of the interpretation process is a representation of the speech act *qua* intentional action: the speaker is construed as (intentionally) communicating that *p*. But to reach that representation only semantic interpretation is necessary, in normal cases. Representations of the speaker's beliefs and intentions play no causal role in *arriving at* the interpretation of the speech act

As I already said, my aim in this paper is to consider what happens to this debate when the standard view is criticized and replaced by a more refined view which pays due regard to the phenomenon of semantic underdetermination. Before turning to that issue, however, I would like to mention a possible inferentialist response to the anti-inferentialist challenge.

The inferentialist may reply as follows. Let us grant, for the sake of the argument, that representations of the speaker's beliefs and intentions play no causal role in arriving at the interpretation of the speech act. Let us grant that semantic interpretation does all the work (in normal cases). Still, it is not enough for the hearer to come to know that the speaker intends to communicate that *p*. Optimally, communication is the process whereby the adressee *comes to know that p* by interpreting the speaker's utterance of a sentence which says that *p*. By semantic interpretation the hearer arrives at the conclusion that the sentence says that *p*, and this

automatically yields an interpretation for the speech act: the speaker is automatically understood as intending to communicate that *p*. But how is the hearer to get to believe that *p*? This can only be inferentially derived. Assuming that the speaker is sincere and well-informed, the hearer can infer that *p* from the fact that the speaker intends to communicate that *p*. So the representation of the speaker's communicative intentions, which results from semantic interpretation, itself plays a causal role in arriving at the representation that *p*, which is the ultimate output of the interpretation process.

That reply will not convince the anti-inferentialist, who will find it question-begging. For the anti-inferentialist, the step from 'the speaker intends to communicate that *p'* to *'p'* is as automatic, as unreflective, as the step from 'the sentence says that *p'* to 'the speaker intends to communicate that *p'*. Normally, the hearer believes what he is told, or at least, he gets the information that *p* when he is told that *p*. Credulity, Reid taught us, is an instinct (Reid 1970: chapter 6, §24). Only when there is something wrong does the hearer suspend or inhibit the automatic transition which characterizes the normal cases of linguistic communication. On the anti-inferentialist view, then, communication is as direct as perception. As Millikan writes, 'the most usual way that people understand the speech of others [is] by translating directly, without inference, into beliefs' (Millikan 1984: 67).[2]

---

[2] For a similar position, see Burge (1993) and McDowell (1980, 1993). As they both point out, the knowledge gained when one is told that *p* in a normal communication situation is no more justified by an inference from the reliability of the source than the knowledge gained through perception is justified by an inference from the reliability of our senses. The reliability of our senses, or the reliability of our informants, is a background condition, just as the proper functioning of memory is a background condition for making long deductions. The proper functioning of memory is not one of the premisses in the deduction, however, and similarly, the reliability of our senses, or the reliability of our informants, is a background condition for the perceptual or communicational acquisition of knowledge, rather than a premiss in a knowledge-providing inference. Of course, doubt is cast on the deliverances of perception or communication whenever there is reason to believe that the background condition is not satisfied. Unless there is such a reason, however, the deliverances of perception and communication count as knowledge

Having disposed of this inferentialist reply to the anti-inferentialist argument, we can turn to what is in effect the most powerful weapon in the inferentialist's hands: the argument from underdetermination (section 4). I will argue that this argument is not conclusive either, appearances notwithstanding. The phenomenon of underdetermination shows that the standard picture must be given up and replaced by a more refined picture in which pragmatic interpretation is unquestionably dominant (section 3). It shows that *semantic interpretation cannot determine the content of the speech act even in normal cases*, because it cannot even determine the truth-conditions of the sentence or 'what the sentence says'. Truth-conditional content is a matter of pragmatics, not semantics. Still, I will argue, this is not sufficient to refute the anti-inferentialist position, which can easily be reinstated in the new framework.

3. Semantic Underdetermination

When a sentence contains a genuine indexical, like 'I' or 'tomorrow', the meaning of the indexical (its character) contextually determines its content in a very straightforward manner. Insofar as the contextual assignment of values to indexicals proceeds according to linguistic rules (e.g. the rule that 'I' refers to the speaker), there is no reason not to consider that aspect of content-determination as part of *semantic* interpretation. The content of an indexical expression is context-dependent, but the type of context-dependence at issue has nothing to do with the radical form of context-dependence which affects *speaker's* meaning and makes it necessary to appeal to pragmatic interpretation. The hallmark of the more radical form of context-dependence (background-dependence, as I call it) is the fact that *any piece of contextual information may be relevant.* But the context that comes into play in the semantic interpretation of indexicals is not the total pragmatic context; it is a very limited context which contains

---

and do not need to be justified by an inference. Such an inference, McDowell notes, would never yield knowledge anyway.

only a few aspects of the pragmatic context: who speaks, when, where, and so forth. As Bach puts it,

> There are two sorts of contextual information, one much more restricted in scope and limited in role than the other. Information that plays the limited role of combining with linguistic information to determine content (in the sense of fixing it) is restricted to a short list of variables, such as the identity of the speaker and the hearer and the time and place of an utterance. Contextual information in the broad sense is anything that the hearer is to take into account to determine (in the sense of ascertain) the speaker's communicative intention. (Bach, 1997: 39)

Thus Bach distinguishes two sorts of context: *wide context* concerns any contextual information relevant to determining the speaker's intention and to the successful and felicitous performance of the speech act; *narrow context* concerns information specifically relevant to determining the semantic values of indexicals and is limited to a short list of contextual variables. Narrow context is semantic, wide context pragmatic.

In terms of this distinction, let me state what I take to be the most serious problem with the standard picture expounded in section 1. The problem is this: most context-sensitive expressions are *semantically underdetermined* rather than indexical in the strict sense. A possessive phrase such as 'John's car' means something like *the car that bears relation R to John*, where 'R' is a free variable. The free variable must be contextually assigned a particular value; but *that value is not determined by a rule and is not a function of a particular aspect of the narrow context*. What a given occurrence of the phrase 'John's car' means ultimately depends upon what the speaker who utters it means. It therefore depends upon the wide context. That is true of all semantically underdetermined expressions: their semantic value varies from occurrence to occurrence, just as the semantic value of indexicals does, yet it varies not as a function of some objective feature of the narrow context but as a function of what the speaker means. It follows that semantic interpretation by itself cannot determine what is said by a sentence containing such an expression: for the semantic value of the expression —

its own contribution to what is said — is a matter of speaker's meaning, and can only be determined by pragmatic interpretation.

Even if we restrict our attention to expressions traditionally classified as indexicals, we see that they involve a good deal of semantic underdetermination. This is true, in particular, of demonstratives. The reference of a demonstrative cannot be determined by a rule, like the rule that 'I' refers to the speaker. It is generally assumed that there is such a rule, namely the rule that the demonstrative refers to the object which happens to be demonstrated or which happens to be the most salient, in the context at hand. But the notions of 'demonstration' and 'salience' are pragmatic notions in disguise. They cannot be cashed out in terms merely of the narrow context. Ultimately, a demonstrative refers to *what the speaker who uses it refers to by using it*.

To be sure, one can make that into a semantic rule. One can say that the character of a demonstrative is the rule that it refers to what the speaker intends to refer to. As a result, one will add to the narrow context a sequence of 'speaker's intended referents', in such a way that the nth demonstrative in the sentence will refer to the nth member of the sequence. Formally that is fine, but philosophically it is clear that one is cheating. We pretend that we can manage with a limited, narrow notion of context of the sort we need for handling indexicals, while in fact we can only determine the speaker's intended referent (hence the semantic referent, which depends upon the speaker's intended referent) by resorting to pragmatic interpretation and relying on the *wide* context.

We encounter the same sort of problem even with expressions like 'here' and 'now' which are traditionally considered as *pure* indexicals (rather than demonstratives). Their semantic value is the time or place of the context respectively. But what counts as the time and place of the context? How inclusive must the time or place in question be? It depends on what the speaker means, hence, again, on the wide context. As is well-known, 'here' can refer to this room, this building, this city, this country, etc., and the same underdetermination affects 'now'. We can maintain that the character of 'here' and 'now' is the rule that the expression refers to 'the' time or 'the' place of the context — a

rule   which automatically determines a content, given a (narrow) context in which the time and place parameters are given specific values; but then we have to let a pragmatic process take place to fix the values in question, that is, to determine *which* narrow context, among indefinitely many candidates compatible with the facts of the utterance, serves as argument to the character function. On the resulting view the (narrow) context with respect to which an utterance is interpreted is not *given*, it is not determined automatically by objective facts like where and when the utterance takes place, but it is determined by the speaker's intention and the wide context. Again, we reach the conclusion that pragmatic interpretation has a role to play in determining the content of the utterance, in such a case.

To sum up, either semantic interpretation delivers something gappy, and pragmatic interpretation must fill the gaps until we reach a complete proposition ; or we run semantic interpretation only after we have used pragmatic interpretation to pre-determine the values of semantically underdetermined expressions, which values we artificially feed into the narrow context. Either way, semantic interpretation by itself is powerless to determine what is said, when the sentence contains a semantically underdetermined expression.

Now I take it that such expressions can be found all over the place. Moreover, semantic underdetermination is not limited to particular lexical items. One can follow Waismann and argue that the satisfaction conditions of any empirical predicate are semantically underdetermined and subject to pragmatic interpretation. There is also constructional underdetermination. For example consider something as simple as the Adjective+Noun construction, as in 'red pen'. What counts as a red pen? A pen that is red. But when does a pen count as red? That depends upon the wide context. The satisfaction conditions of 'red pen' can only be determined by pragmatic interpretation. (Of course there are default interpretations; but they can be overriden and I think that establishes their essentially pragmatic nature.)

4. Does Underdetermination Support Inferentialism?

It seems that semantic underdetermination provides a knock-down argument in favour of inferentialism. What semantic underdetermination shows is this. Even in normal cases, one can determine the content of the speech act only by appealing to pragmatic considerations concerning what the speaker means, what his intentions are. For example, we must know who the speaker intends to refer to when he says 'that guy' if we are to understand the content of the utterance in which the phrase occurs. Even when the speaker speaks literally and 'means what he says', interpreting an utterance requires consideration of the speaker's beliefs and intentions.

But the situation is more complex. To begin with, when we say that in order to understand an utterance with the phrase 'that guy' we must 'know who the speaker intends to refer to', this is ambiguous. What is true is that to understand what is said we must identify the referent. We must think of the right person as being 'that guy'. And the right person *is* the person the speaker intends to refer to. But we need not think of the person in question *as* the person the speaker intends to refer to. More generally, we don't have to think about the speaker's communicative intentions; we merely have to grasp the right meaning, where the right meaning *is* the speaker's meaning — but we don't have to think of it *as* the speaker's meaning.

There are psychological processes in us that take us from the meaning of the sentence to the content of the utterance (corresponding to the speaker's intent). Those processes are pragmatic, not semantic. They exhibit the distinguishing characteristic of pragmatic processes: the interpretation they yield is potentially sensitive to any change in the (wide) context. But the processes in question, *qua* causal processes somehow realized in the brain, need not involve the representation of the speaker's beliefs and intentions,[3] and they need not be inferential (rather than, say, associative). They yield an output which passes muster only if it compatible with what we know of the speaker

---

[3] I do not mean to deny that the processes in question *may* involve the representation of the speaker's intentions. See footnote 6 below.

as a rational agent, but consideration of the speaker's rationality and his web of beliefs and desires comes into the picture only if there is something wrong with the output of the automatic pragmatic processes, whatever they are, that deliver the prima facie interpretation for the utterance.

It follows that semantic underdetermination is compatible with anti-inferentialism. We can accept that semantic interpretation cannot deliver the content of the speech act, even in normal cases; pragmatic processes are involved in all cases in the determination of truth-conditional content. But those pragmatic processes that are involved in the determination of truth-conditional content — *primary* pragmatic processes, in my terminology — need not involve an inference from premises concerning what the speaker can possibly intend by his utterance. *Indeed, they need not involve any inference at all*.

Primary pragmatic processes are very different from secondary pragmatic processes (Recanati 1993). Secondary pragmatic processes are those that come into play to determine what the speaker means on the basis of what he says when what the speaker means goes beyond what he says — when he 'implies' various things by saying what he says. Conversational implicatures, in the strict sense, are inferentially derived from premises concerning the speaker's intentions in saying what he says. They are arrived at by answering questions such as, 'Why is the speaker saying what he says?' Secondary pragmatic processes are background-dependent, like primary pragmatic processes, but they have further properties which distinguish them from primary pragmatic processes. First, the interpreter has to be aware of what is said, aware of what is implied, and aware of the inferential connection between them. (This property I call 'availability'.) Second, the interpreter's representation of the speaker's underlying intentions plays a causal role in the mental process of determining what the speaker implies. As Grice emphasized, implicature-determination in the strict sense is a reflective process. Instead of merely retrieving what is said through the operation of unconscious, primary pragmatic processes, we reflect on the fact that the speaker says what he says and use that fact, together with background knowledge, to infer what the

speaker means without saying it. As Millikan writes, 'the true communicator is in a position to tinker with the mechanisms of normal language flow, is sensitive to symptoms that the other is tinkering with these mechanisms, and can rise above these automatic mechanisms if necessary' (Millikan, 1984: 69). That is what happens in *special* cases. The retrieving of conversational implicatures, in particular, involves reflective capacities that are not exercised in what Millikan calls 'normal language flow'.

The contrast I am trying to draw can also be stated in terms of different levels of processing. The determination of what is said takes place at a sub-personal level, much as the determination of what we see. But the determination of what the speaker implies takes place at the personal level, much like the determination of the consequences of what we see. (Seeing John's car, I infer that he did not leave.) The crucial fact is that pragmatic, background-dependent processes may well take place at a sub-personal level in an automatic and non-reflective manner. Such processes are not 'inferential' in the strong sense in which secondary pragmatic processes are inferential.

5. Two Senses of 'Inferential'

Theorists of pragmatics like Bach and Harnish, Sperber and Wilson and others in the Gricean tradition insist that linguistic communication is 'inferential'. What exactly do they mean?

In introducing the idea, Sperber and Wilson observe that

> When... [ordinary hearers] are asked to explain how they know which interpretation is correct, they generally offer something that looks like a truncated logical argument: the speaker must have intended this interpretation rather than that, because this is the only interpretation that is true; or the only one that gives the

required information; or the only one that makes sense. (Sperber and Wilson, 1986: 13)[4]

As Millikan points out, such inferences are made only when the language users rise above the automatic mechanisms of normal language flow. There is no doubt that our ability to do so is an essential part of our conversational competence; no one denies this. But the question at issue is: *How essential* is that meta-discursive ability? Can linguistic communication proceed without meta-discursive inference, at some basic level, or is it from the very start constituted by it?

The anti-inferentialists claim that the sort of meta-discursive ability illustrated in the above quotation from Sperber and Wilson is not constitutive of linguistic communication. As Burge writes,

We seem normally to understand content in a way whose unconscious details (…) are not accessible via ordinary reflection. To be entitled to believe what one is told, one need not understand or be able to justify any transition from perceptual beliefs about words to understanding of and belief in the words' content. One can, of course, come to understand certain inferences from words to contents. Such empirical meta-skills do enrich communication. But they are not indispensable to it. To be justified in understanding, we have to reason empirically about what we perceive only when communication runs into trouble, or when special, contextual, nonliteral expressive devices are used. (Burge, 1993: 477)

There is a sense, then, in which (if the foregoing is true) linguistic communication is *not* inferential at the most basic level. But is it the same sense in which theorists of pragmatics claim that linguistic communication is inferential? Presumably not. Communication is non-inferential in the sense that it does not rest on reasoning. Reid defines 'reasoning' as 'the process by which we pass from one judgment

---

[4] As we shall see, these intuitive meta-discursive reconstructions are not endorsed by Sperber and Wilson. They appeal to 'a form of unconscious inference simpler than conscious reasoning' (Sperber and Wilson 1987 : 736).

to another which is the consequence of it. Accordingly, our judgments are distinguished into intuitive, which are not grounded upon any preceding judgment, and discursive, which are deduced from some preceding judgment by reasoning' (Reid, 1969: 710). Evidently, comprehension is intuitive rather than discursive, except in the special situations in which one adopts a reflective stance towards the ongoing discourse. Normally we do not have to reason to understand what the others are saying: the judgment that the speaker has said that p is made directly upon hearing the utterance, without being inferentially grounded in some prior judgment to the effect that the speaker has uttered sentence S. But the inferentialists do not deny that comprehension is intuitive rather than discursive. When they say that communication is inferential, they mean it in a sense which is compatible with its being intuitive. According to them, there are two sorts of inference (or two sorts of 'reasoning'). One is conscious, explicit inference — what Reid calls 'reasoning'. Let us call that 'inference in the narrow sense'. But there is also inference in the broad sense: a type of inference which may well occur unconsciously, in such a way that the judging subject is aware only of the conclusion of the inference (which is, therefore, available not as the conclusion of an inference but as an immediate, intuitive judgment).

The inferentialist admits, indeed emphasizes, that the inferential procedure underlying ordinary understanding is unconscious. In special cases the inference is made explicit at a reflective level; but inference is there all the time, they maintain, whether or not it is available to consciousness.   As Sperber says, 'when most of us talk of reasoning, we think of an occasional, conscious, difficult, and rather slow mental activity. What modern psychology has shown is that something like reasoning goes on all the time — unconsciously, painlessly, and fast.' (Sperber, 1995: 195) It is in that modern psychological sense that communication is said to be 'inferential'.

That there is such a broad notion of inference in contemporary cognitive science is beyond question. For example, Marr writes that 'the true heart of visual perception is the inference from the structure of an image about the structure of the real world outside' (Marr 1982: 68). Insofar as this is the notion of inference appealed to by the

inferentialists, the debate turns out to be verbal. Burge himself, in the first sentence of the passage I have quoted, grants that unconscious, sub-personal 'inferences' may well take place in the process of understanding. He writes : 'We seem normally to understand content in a way whose unconscious details *(inferential or otherwise)* are not accessible via ordinary reflection.' (Burge, 1993 : 477 ; italics mine.) This is compatible with communication's not being inferential in the narrow sense — a sense that *requires* conscious accessibility.

In view of the existence of the two senses, it is misleading to argue against the anti-inferentialist position, which says that communication is not inferential in the narrow sense, on the grounds that communication is inferential in the broad sense. But — if I read her correctly — that is what Anne Bezuidenhout does in her critique of Burge (Bezuidenhout 1998). Burge (1993) distinguishes preservative processes like memory and justificatory processes like inference (in the narrow sense).[5]   He claims that 'in interlocution, perception of utterances makes possible the passage of propositional content from one mind to another rather as purely preservative memory makes possible the preservation of propositional content from one time to another' (Burge, 1993: 481). The cognitive processes underlying comprehension are therefore preservative rather than justificatory. Bezuidenhout's argument against Burge has the following form:

(i) Burge grants that 'inferential processes do play a role in justification' (Bezuidenhout, 1998: 269).

(ii) But, as Sperber and Wilson and others have shown, the cognitive processes underlying communication are inferential.

(iii) It follows that Burge is wrong: 'the underlying processes do play a justificatory role. This is because the underlying processes are essentially inferential, and it has already been established... that inferential processes play a justificatory role.' (Bezuidenhout, 1998: 270)

---

[5] See footnote 2.

That argument, it seems to me, trades on a confusion between the two senses of 'inference'. Burge grants that inference is justificatory only to the extent that 'inference' is understood in the narrow sense. But in (ii), 'inference' had better be taken in the broad sense (since the inferentialist position defended by Bezuidenhout would be indefensible if 'inference' were taken in the narrow sense).

6. Spontaneous Inference and the Availability Condition

What I call inference in the narrow sense corresponds to Reid's description of 'reasoning' understood as a conscious activity. I claim that, in communication, such inferences take place in *special* cases. My distinction between primary and secondary pragmatic processes rests on the fact that the latter are, while the former are not, inferential in that narrow sense. But the notion of conscious inference that is relevant here does not quite correspond to Sperber's description. Sperber contrasts explicit, conscious reasoning, which is a voluntary, effortful and slow activity, with spontanous inference, which is effortless, fast, and takes place unconsciously. That distinction I find misleading in the present context. Like Sperber, I think that inference in the narrow sense is essentially conscious: it takes place only if one judgment (the conclusion) is grounded in another judgment (the premise), and if both judgments, as well as the fact that one is grounded in and justified by the other, are available (consciously accessible) to the judging subject. Secondary pragmatic processes are inferential in the narrow sense because they satisfy that essential condition — the availability condition. But it would be wrong to claim that narrow inferences are necessarily effortful, slow and under voluntary control. Among inferences in the narrow sense, some — including those that underly the retrieval of conversational implicatures — are typically spontaneous: the inference is drawn more or less automatically.

Consider Sperber's following examples of spontanous inference: you hear the doorbell ringing, and you form the belief 'There is someone at the door' (Sperber, 1997:

77-78). That belief is inferentially derived from a prior belief (to the effect that the doorbell is ringing) directly based on perception . 'If challenged, you might be able to produce, *ex post facto*, a missing premise that, together with the perceptual belief, warrants the inferential belief. However, the fact is that you arrived at the inferential belief without engaging in deliberate or conscious inference' (Sperber, 1997: 78). But in all the cases of spontaneous inference mentioned by Sperber in the passage from which I've just quoted, including the doorbell example, the availability condition *is* satisfied: the subject makes two judgments, one based on perception, the other based on inference from the first belief. The two judgments are conscious and available to the subject. Moreover, the subject is aware that the second judgment is grounded in, and justified by, the first. If she says, 'There is someone at the door', and is asked 'How do you know?', she will reply something like: 'The doorbell is ringing'. The perceptual judgment to the effect that the doorbell is ringing justifies, and is known to justify, the nonperceptual judgment that there is someone at the door. Since there are two judgments standing in the appropriate relation to each other, Reid's definition of 'reasoning' applies, even though this piece of reasoning is spontaneous, effortless and fast. The only thing that is not conscious here are the inferential steps needed to bridge the gap between the first, perceptual belief and the second, inferential belief. Still, the inference, though not *explicit*, is *conscious* in the sense that the availability condition is satisfied. This is in contrast to cases in which the availability condition is *not* satisfied: cases, for example, where the subject is aware only of one judgment, the alleged inferential source of that judgment being unavailable to consciousness; or cases in which both judgments are available, but the subject is unaware of one being inferentially derived from the other.

I conclude that there are two sorts of inference in the narrow sense : explicit reasoning, and spontaneous inference. In both cases the availability condition is satisfied. Inferences in the broad sense are characterized by the fact that the availability condition is *not* satisfied. Now it is my contention that, if primary pragmatic processes are indeed inferential, they are inferential only in the broad sense. The interpreter is not

aware that his judgment, to the effect that the speaker has said that *p*, is inferentially derived from a prior judgment (e.g. a judgment to the effect that the speaker has uttered sentence S). Similarly, some theorists claim that perceptual judgments themselves are inferential, but the inferences at issue do not satisfy the availability condition: the perceiver is not aware that his perceptual judgment that *p* is based on a prior judgment (to the effect that things seem thus and so, or whatever).

If the subject says 'The doorbell is ringing' and is asked 'How do you know?', she will reply: 'I can hear it'. That reply merely indicates the perceptual nature of the current judgment 'The doorbell is ringing'; it does not reveal a prior judgment on which it is based. The current judgment is not based on another judgment, as in reasoning, it is directly based on perception and is therefore justified *qua* perceptual judgment, without any need for a grounding premiss. The same thing holds for communication. If asked 'How do you know', the subject may reply: 'John told me'. Here the response indicates the communicational nature of the judgment at stake. If I am told that the doorbell is ringing, my judgment that it is ringing is justified unless there are reasons to disbelieve what I am told, just as, if I hear the doorbell ringing, my judgment that it is ringing is justified unless there are reasons to disbelieve the testimony of my senses. This is not to deny that, perhaps, an inference is involved in forming the perceptual or the communicational judgment. But that inference, if it takes place, does not satisfy the availability condition: it is not an inference *for the subject*, and it plays no justificatory role in her mental life.[6]

---

[6] Once it is clear that we are dealing with sub-personal microprocesses rather than with genuine thought processes, we can reconsider the issue of speaker's intentions and their role in ordinary understanding. It is plausible that, among the sub-personal inferences which sustain ordinary understanding, some take a representation of the speaker's intentions among their inputs. For example, Paul Bloom (in his contribution to the Oxford workshop) reminded us that in language acquisition, the gaze of the speaker and other indicators of the speaker's intention play a crucial role in reference-fixing. The same cues also play a role in ordinary comprehension (e.g. in the interpretation of demonstratives). That is compatible with what I said, insofar as the relevant mechanisms of intention-detection are recruited as part of the

7. An Intermediate Sort of Inference ?

Since there are (at least) two senses of 'inference', the inferentialist/anti-inferentialist debate turns out to be verbal to a large extent. The anti-inferentialist does not deny that communication, or even perception for that matter, may be inferential in the broad sense; nor does the inferentialist deny that communication, like perception, is noninferential in the narrow sense. Still, I think there is room for genuine disagreement, on several counts.

First, the anti-inferentialist emphasizes the difference between inference in the narrow sense (inference at the personal level) and inference in the broad sense (inference at the sub-personal level), while the inferentialist tends to downplay it. To be sure, the inferentialist (e.g. Sperber) explicitly distinguishes between inferences in the ordinary sense and (possibly unconscious) inferences in the psychological sense. But (i) he suggests that the same type of process occurs in both cases, accompanied or unaccompanied by a state of conscious awareness; and (ii) by mislocating the distinction and phrasing it in terms of explicitness and voluntary control he neglects the crucial divide between cases in which the availability condition is satisfied (including spontaneous inference) and cases in which it is not (e.g. communicational or perceptual inference). That divide is, for the anti-inferentialist, fundamental. As Bezuidenhout notes (1998 : 268), Burge suggests that mental processes are justificatorily relevant

---

sub-personal machinery underlying comprehension. To be sure, the mechanisms in question unambiguously belong to the mind-reading faculty, but their relation to that faculty must be properly understood. As Baron-Cohen pointed out, the mechanisms of intention-detection which play a role in normal comprehension (e.g. gaze tracking and the shared attention mechanism) are fundamental to the emergence of a full-fledged metarepresentational capacity of the sort that is exercised when we think about the mental states of others (Baron Cohen 1995: 30-58). That complex capacity *presupposes* the primitive mechanisms of gaze-tracking and intention-detection — it builds upon them — but they do not, themselves, presuppose the exercise of that capacity.

only if they are accessible via ordinary reflection, that is, only when the availability condition is satisfied.

Second and more important, the anti-inferentialist stresses that communication is exactly like perception: in both cases information is acquired 'directly', i.e. without inference (in the narrow sense). That view goes back to Thomas Reid.[7] Inferentialists, on the other hand, claim that *communication is inferential in a sense in which perception is not.* This is obvious in Sperber's treatment of these matters. Communication is said to rest on spontaneous inference, and spontaneous inference is (rightly) contrasted with perception: I *hear* the doorbell ringing, but I *infer* that there is someone at the door. If I am told that *p*, this is again a case of inference, according to Sperber: I *hear* the words, but I *infer* that the speaker is telling me that *p*. Communication is on the same side as perception for the anti-inferentialist, with inference (in the narrow sense) standing on the other side of the divide. For the inferentialist, communication stands on the same side as spontaneous inference, in contrast to perception which stands on the other side.

If we use the availability condition as the criterion for classifying inferences, we are bound to disagree with Sperber. Spontaneous inferences of the sort he describes satisfy the availability condition, while primary pragmatic inferences do not (section 6). In this respect primary pragmatic inferences are exactly like the 'inferences' involved in perception :   the anti-inferentialist viewpoint is therefore vindicated. But perhaps availability is not the right (or the only) criterion to use. Indeed Sperber and Wilson, who defend a variety of inferentialism, use a different criterion. They, too, distinguish inferences in the broad sense from a more restricted subclass, but the subclass they focus on is delimited not in terms of conscious availability, but in terms of their conceptual nature.

Marr uses 'inference' in a *very* broad sense. Nothing prevents the input to an inference in that sense (an M-inference, for short) from being nonconceptual.

---

[7] See Reid, 1970. The view is endorsed by many contemporary writers, including Evans, 1982 and Fodor, 1983.

Representations of intensity changes in terms of zero-crossings are clearly nonconceptual (Crane, 1990), yet Marr describes the transition from such proximal representations to distal representations of edges as an inference. Fodor (1983) uses 'inference' in the same   broad way. But Sperber and Wilson use 'inference' much more restrictively.   For there to be inference, they claim, the transition must be from a conceptual representation to a conceptual representation. Moroever, the transition must be truth-preserving. Only if these conditions are satisfied will a cognitive transition count as a genuine 'inference'.[8]

According to Sperber and Wilson, primary pragmatic inferences are genuine inferences by this criterion, whereas perceptual inferences are not. The inferentialist viewpoint is therefore vindicated : since inferences in the narrow sense (N-inferences, for short) are themselves truth-preserving operations on conceptual representations, there is a sense in which primary pragmatic inferences are more similar to N-inferences (despite the noted difference in terms of availability) than to M-inferences and other sub-personal computations.

In Sperber and Wilson's framework, primary pragmatic inferences turn out to be *intermediate* between N-inferences and M-inferences. The availability condition is not satisfied — the inference is unconscious — yet we are still dealing with conceptual representations manipulated under constraints of rationality. Note that, like N-inferences, primary pragmatic inferences are background-sensitive and display a high level of cognitive penetrability. That feature is supposed to distinguish the 'central systems' of the human mind, which manipulate conceptual representations under constraints of rationality, from 'input systems' like vision, which are modular and 'informationally encapsulated' (Fodor, 1983). As Crane puts it,

The inferences a thinker is disposed to make (...) are constrained only by rationality, which allows the mind to range over its whole territory for its material. The

---

[8]   '[Whenever] the input and the output of a [computation] are not related as premise and conclusion in an argument, the computation is not interpretable as inferential' (Sperber and Wilson 1987 : 137).

> resources of the visual system are, by contrast, severely restricted. Although we can treat it as deducing consequences from premises, the contents of these premises are not holistically related in the way the contents of beliefs are. (Crane, 1990: 156)

In this respect, primary pragmatic inferences turn out to belong to the same realm as inferences in the narrow sense. They display the holistic properties that characterize operations of the central systems (Wilson and Sperber, 1986: 584-585, Bezuidenhout, 1998: 271-272).

At this point, we might conclude that the inferentialist and the anti-inferentialist viewpoints are both vindicated. From a certain perspective, communication is like perception and unlike inference (in the narrow sense). From another perspective, just the opposite is the case : communication is like reasoning and unlike perception. On this ecumenical view, inferentialism and anti-inferentialism correspond to the use of two distinct, equally legitimate criteria : the conscious availability of the inferences, and the conceptual nature of the representations operated on. —I will refrain from so concluding, however, because I remain unconvinced by the argument I have just presented on behalf of Sperber and Wilson. For reasons I am going to spell out, I am not convinced that primary pragmatic inferences are truth-preserving operations on conceptual representations ; and I am not convinced that cognitive penetrability can be invoked to contrast communication and perception.

8. 'Conceptual' Inferences in Communication and Perception

According to Sperber and Wilson, primary pragmatic processes are 'inferential' in the (intermediate) sense that they are truth-preserving operations on conceptual representations. Are they really ?

One of the primary pragmatic processes involved in the determination of what is said consists in 'saturating' or 'completing' a schematic meaning in order to yield a truth-evaluable representation. All the inferentialists rightly emphasize the abstract, schematic character of the 'linguistic meanings' which serve as input to primary

pragmatic processes. That schematic, abstract character renders somewhat dubious the claim that those meanings (which,   according to Sperber and Wilson, never surface to consciousness) are themselves 'conceptual representations'. I would rather say that they are nonconceptual, like the representation of intensity changes in terms of zero-crossings, or at least not *fully* conceptual.

Admittedly,   there are primary pragmatic processes which, unlike saturation, take fully conceptual material as input. Thus free enrichment maps a conceptual representation to another, more specific conceptual representation. The primary pragmatic process which I dubbed 'transfer' also maps conceptual representations to conceptual representations. But in all these cases it is dubious that the process is truth-preserving. To begin with, the process is local — it operates at the level of constituents rather than the level of complete, truth-evaluable representations (Recanati, 1993, 1995). Even if we disregard this feature and construe the process as a global mapping from propositional representations to propositional representations, still it is obviously *not* truth-preserving.   In the case of transfer, the relation between input and output is associative rather than logical. In the case of free enrichment, there is a logical relation between the input and the output : the latter entails the former. This does not make free enrichment truth-preserving, however, because the entailment relation goes in the wrong direction.

The only way I can make sense of the claim that primary pragmatic processes are truth-preserving operations on conceptual representations is by moving to the meta-discursive level: there we find a transition (from *the speaker utters sentence S* to *the speaker says that* p) which is arguably truth-preserving. The input to that inference is a description of the speaker's utterance of a sentence possessing a certain meaning, and the output is a description of the speaker's locutionary act. The extra premisses used in the inference correspond to the contextual facts which have a bearing on the utterance's truth-conditions. But, as Sperber and Wilson themselves pointed out (in response to Philip Pettit, whose argument anticipates that of this paper), we can, by moving to the

meta-discursive level, make even the transition from phonetic representation to semantic representation (what Sperber and Wilson call 'decoding') truth-preserving :

> It is true [as Pettit claims in his critique of the decoding/inference contrast] that linguistic input processes [themselves] can be characterised as inferential… From the premise that Mary uttered sounds with a certain phonetic representation, Peter may be able to infer that she said she was leaving. But this does not affect the point at issue : The linguistic decoding processes embedded in this broader inferential process are not themselves inferential, in the sense that they are not truth-preserving (…) operations on conceptual representations (Sperber and Wilson, 1987 : 737).

I think exactly the same considerations apply to primary pragmatic processes: they are not *themselves* inferential (in the sense intended by Sperber and Wilson), even if they can be *embedded in* a broader inferential process by moving to the meta-discursive level.

To sum up, it is not obvious that primary pragmatic processes operate on conceptual representations, and, when they do so, it is not obvious that they are truth-preserving. Still, I agree that they are cognitively penetrable and display the property of background-dependence — arguably the hallmark of conceptual processes. Perhaps this speaks in favour of the 'intermediate' level of inference posited by Sperber and Wilson. Be that as it may, I do not think this feature supports the inferentialist's claim that communication is inferential (at the relevant level) while perception is not.

As Oaksford and Chater (and many others) pointed out, knowledge-rich defeasible inference processes are 'implicated throughout almost every area of cognitive activity' (Oaksford and Chater, 1991: 10). No less than communication, perception is a case in point. What we see (in the 'cognitive' sense, involving identification of the perceived object) is highly context-sensitive (Dretske, 1990). Searle claims that the content of our perceptions holistically depends upon 'the

background' just as the content of our utterances depends upon the background.[9] He gives the following example:

> Suppose I am standing in front of a house looking at it; in so doing I will have a certain visual experience with a certain Intentional content, i.e. certain conditions of satisfaction; but suppose now as part of the background assumptions I assume I am on a Hollywood movie set and all of the buildings are just papier maché façades. This assumption would not only give us different conditions of satisfaction; it would even alter the way the façade of the house looks to us, in the same way that the sentence 'Cut the grass!' would be interpreted differently if we thought that the background was such that we were supposed to slice the grass rather than mow it. (Searle, 1980: 231)

> It is part of the content of my visual experience when I look at a whole house that I *expect* the rest of the house to be there if, for example, I enter the house or go around to the back. In these sorts of cases the character of the visual experience and its conditions of satisfaction will be affected by the content of the beliefs that one has about the perceptual situation. I am not going beyond the content of my visual experience when I say, 'I see a house' instead of 'I see the façade of a house', for, though the optical stimuli may be the same, the conditions of satisfaction in the former case are that there should be a whole house there. (Searle, 1983: 54-55)

This example shows how 'inferential' our perception can be. Of course, what I see does not result from an inference in the *narrow* sense, as when I infer John's presence from the perception of his car. It results from an inference in the broad sense — an inference that takes place at the sub-personal level. Yet that inference too displays the relevant feature of holistic background-dependence.

Another example. Following Sperber, I contrasted the perceptual belief that the doorbell is ringing with the inferential belief that there is someone at the door. Now the perceptual belief that the doorbell is ringing itself is 'inferential', since it involves *identifying* the sound as that of the doorbell ringing. The difference between the two beliefs is that one is inferential in the narrow sense, while the other is inferential in the

---

[9] For a critical discussion of Searle's ideas on this topic, see Recanati forthcoming.

broad sense : the inference at issue takes place at the sub-personal level. Still, *the inference involved in the latter case* (the perceptual belief) *is no less defeasible and background-dependent than the 'narrow' inference that there is someone at the door*.

What all this shows is that, within the sub-personal level, there is an upper layer which is cognitively penetrable (see Dretske 1990 for a similar point). Just as we distinguished explicit reasoning and spontaneous inference as two sub-classes within the class of inferences in the narrow sense, we must distinguish two sub-classes within the class of inferences in the broad sense : those that are cognitively penetrable and background-dependent (e.g. the inferences involved in perceptual identification), and those that are not. If that is right, then, if there is an intermediate level of inference which is conceptual (like N-inferences) but unconscious (like M-inferences), as Sperber and Wilson claim, there is no reason to deny that it is involved in perception as much as in communication. Again, we reach the anti-inferentialist conclusion that communication is 'as direct as' perception.

References

Bach, K. 1997 : The semantics-pragmatics distinction: what it is and why it matters. *Linguistiche Berichte,* 8, 33-50.

Bach, K. and Harnish, M. 1979 : *Linguistic Communication and Speech Acts*. Cambridge, MA: MIT Press.

Baron-Cohen, S. 1995 : Mindblindness: An Essay on Autism and Theory of Mind. Cambridge, MA : MIT Press.

Bezuidenhout, A. 1998 : Is verbal communication a purely preservative process? *Philosophical Review*, 107, 261-288.

Burge, T. 1993 : Content preservation. *Philosophical Review,* 102, 457-488.

Crane, T. 1990 : The nonconceptual content of experience. In T. Crane (ed.), *The Contents of Experience*. Cambridge: Cambridge University Press.

Dretske, F. 1990 : Seeing, believing, and knowing. In D. Osherson et al. (eds.), *An Invitation to Cognitive Science,* vol. 3. Cambridge, MA: MIT Press.

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

Fodor, J. 1983 : *The Modularity of Mind*. Cambridge, MA: MIT Press.

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

Marr, D. 1968 : *Vision: A computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: W.H. Freeman and C°.

McDowell, J. 1980 : Meaning, Communication, and Knowledge. In Z. van Straaten (ed.), *Philosophical Subjects: Essays Presented to P.F. Strawson*. Oxford: Clarendon Press.

McDowell, J. 1993 : Knowledge by Hearsay. In B. Matilal and A. Chakrabarti (eds.), *Knowing from Words: Western and Indian Philosophical Analysis of Understanding and Testimony*. Dordrecht: Kluwer.

Millikan, R. 1984 : *Language, Thought, and Other Biological Categories*. Cambridge, MA: MIT Press.

Oaksford, M. and Chater, N. 1991 : Against logicist cognitive science. *Mind and Language,* 6, 1-38.

Pettit, P. 1987 : Inference and information. *Behavioral and Brain Sciences*, 10, 727-729.

Recanati, F. 1993 : *Direct Reference: From Language to Thought*. Oxford: Blackwell.

Recanati, F. 1995 : The alleged priority of literal interpretation. *Cognitive Science,* 19, 207-232.

Recanati, F. forthcoming : The limits of expressibility. In B. Smith (ed.), *John Searle*. Cambridge: Cambridge University Press.

Reid, T. 1969 : *Essays on the Intellectual Powers of Man*. Cambrige, Mass: MIT Press.

Reid, T. 1970 : *An Inquiry into the Human Mind*. Chicago: University of Chicago Press.

Searle, J. 1980 : The background of meaning. In J. Searle, F. Kiefer and M. Bierwisch (eds.), *Speech Act Theory and Pragmatics*. Dordrecht: Reidel.

Searle, J. 1983 : *Intentionality*. Cambridge: Cambridge University Press.

Sperber, D. 1994 : Understanding verbal understanding. In J. Khalfa (ed.), *What is Intelligence?* Cambridge: Cambridge University Press.

Sperber, D. 1995 : How do we communicate? In J. Brockman and K. Matson (eds), *How Things Are: A Science Toolkit to the Mind.* New York: Morrow.

Sperber, D. 1997 : Intuitive and reflective beliefs. *Mind and Language,* 12, 67-83.

Sperber, D. and Wilson, D. 1986 : *Relevance: Communication and Cognition.* Oxford: Blackwell.

Sperber, D. and Wilson, D. 1987 : Précis of *Relevance : Communication and Cognition. Behavioral and Brain Sciences*, 10, 697-754

Waismann, F. 1951 : Verifiability. In A. Flew (ed.), *Logic and Language*, 1st series. Oxford: Blackwell.

Wilson, D. and Sperber, D. 1986 : Pragmatics and modularity. In A. Farley, P. Farley, and K.-E.McCullough (eds.), *The Chicago Linguistic Parassession on Pragmatics and Grammatical Theory*; reprinted in S. Davis (ed.), *Pragmatics: A Reader*. Oxford: Oxford University Press, 1991.

Institut Jean Nicod

CNRS, Paris

[word count : 10037]