Manipulation and cognitive pragmatics: Preliminary hypotheses

Louis de Saussure

University of Neuchâtel

1. Introduction[*]

The paper presents a set of pragmatic hypotheses on the topic of manipulative and ideological discourse; these hypotheses, which are to be further explored and empirically validated, aim at grounding a research project. The general framework of this research is a mechanistic and naturalistic cognitive theory of human communication, Sperber and Wilson's Relevance theory (Sperber and Wilson 1995).[1]

In this introduction, after briefly defending the legitimacy of a cognitive pragmatic approach of manipulation, I consider the main contributions to the problem in that framework.

In section 2, I propose a working definition for *manipulative discourse*, where manipulative discourse is *truth-conditionally / truth-functionally defective*, *doubtful* (when the propositions conveyed are about a state of affairs), or *unacceptable within a given culture* (when the propositions conveyed are about moral issues). Section 3 discusses the typology of strategies involved in discursive manipulation, and notes a commonality between all the strategies; they lead the hearer to problems of understanding, i.e., problems in the process of retrieving a clear informative intention on the part of the speaker. In section 4, I suggest that a central mechanism of manipulation through discursive strategies is organised in a twofold process, managed by the speaker: causing trouble in the hearer's understanding procedure and offering ready-made resolutions of that trouble. I call that mechanism the *trouble-and-resolution* device. This process disturbs the normal processing of intention recovery,

permitting to obtain sincere consent to a proposition regardless of defective arguments, given a particular context. In section 5, I discuss the manipulative discourse with regard to the theory of mind.

Considering a problem like manipulative discourse through the question of *understanding natural language* is quite different from what is generally proposed on the topic within discourse analysis and the social sciences. Many trends defend the view that anything manipulative is best understood with informal tools such as those provided by text-linguistics (see the important works of K. Ehlich in Germany and of J.-M. Adam in French for example, and trends in the Critical Discourse Analysis approach). However, if assuming that manipulation in discourse is primarily achieved during the very construction of meaning in context (which is a first and necessary step for adding beliefs in one's knowledge base), it becomes clearer that theories provide descriptions for that cognitive process, which goes on when exposed to speech or text, are likely to enlighten significantly – and with less intuitive notions – research on manipulation. I suggest, moreover, that such an approach is likely to contribute to the transdisciplinary task of better grounding the fine-grained descriptions done in informal approaches.

A related point is that scholars who favour a social approach to discourse and manipulation often tend to avoid an objectivist standpoint, for various reasons. The most obvious of these reasons is the fear that the analyst can be misled, taking his own subjectivity for objectivity (see Kienpointner in this volume for a discussion). The typical distortion is that the manipulative nature of a discourse would be identified 'at first sight' and intuitively by the analyst (notably because of stereotypes and prejudices), so that he inevitably finds back the manipulative nature of the considered text by means of his own tools of description. This would be, of course, circular and methodologically unacceptable.

But when rejecting any possibility of objectivist, truth-conditional and realist assumptions, the risk of 'reading into' the text one's own expectations may be even greater with more intuitive tools: virtually any text or speech could be considered 'manipulative'.

Scholars in such trends are, of course, conscious of this kind of problem and they generally prefer not to talk of manipulation at all, or they begin by saying that all texts are manipulative but some are *more* manipulative than others.

Now, the same range of objections can be raised against contextual semantic analysis in general. Given an utterance and a context, the analyst knows, because of his 'linguistic

intuition', what the basic meaning of that utterance is. So whatever kind of analysis he would do, he would find out that very meaning he came to intuitively.

When considering semantic and pragmatic analysis of this kind, it becomes clearer that the objection of circularity either applies to any language-related analysis or is false for some reason. I argue for the latter.

The fact is that the analyst who has access to a text can differentiate between what is conveyed by the linguistic strings and what information is needed in the context to build up a consistent and relevant meaning. He does so when analysing ordinary utterances. He does so when noticing, thanks to that very differentiation, that a discourse should be understood under normal circumstances in a way in which it is common knowledge that it was not. Then, the question of intention comes up and, in particular, that of non-benevolent hidden intention.

Therefore I assume that, when comes the time for the analysis of a given manipulative discourse, the ordinary method of the pragmaticist remains unchanged, but addresses deceptive and non-benevolent communication instead of ordinary and benevolent communication; he needs then not to talk only about *understanding* but about the complex way that leads from comprehension to believing.

1.1. Relevant contributions to the issue within the cognitive framework

Cognitive pragmatics has addressed manipulative discourse only recently and to a limited extent; however the few hypotheses that have been explored within this paradigm present a rich ground for further developments.

Choi, Nisbett and Smith (1997) argue that socio-cultural factors concerning information and category salience directly affect human stereotypic reasoning. This study opens a link between changes in salience of information and inferential results which could possibly be exploited manipulatively when building up generalizations and other inductive conclusions.

In Johnson, Grazioli, Jamal and Berryman (2001), an experimental study addresses the problem of deception-detecting in groups of individuals. Although not directly connected with political manipulation, this work is in fact a continuation of Dennett's hypotheses on what he calls the *intentional stance strategy*, where detected inconsistencies in discourse are interpreted in the light of the deceiver's goals and possible actions (Dennett 1989). This

aspect of research is of great interest since, as we recall later, manipulation in discourse is often discovered on the basis of inconsistencies and (formal and non-formal) fallacies.

Regarding political manipulation proper, a few references are available within Relevance theory: Sperber and Wilson (1995), Allott (2002), Allott (this volume), Allott and Rubio Fernandez (2002), Blass (2002) and Carston (2002) all address the problem of terminological misuse with regard to interpretative processes (misuse of concepts and *shallow processing*). They show that often, lexical items are not properly decoded; they just fill a position until some concept can be substituted. A manipulation can take place when some inappropriate word is used to refer to the awaited concept, leading to problematic inferences. For reasons of processing economy, the hearer either provides the proper concept (as in the joke "where did they bury the survivors?" where *survivors* is misread as *dead*), or accepts the item's connotation in situations where the item should be rejected (connotations are part of the *encyclopaedic entry* of concepts, according to Relevance theory), because of some conceptual variation of the item in context. For example the word *democracy*, which is normally understood as *political system where the people take a crucial part in decisions*, is sometimes understood as a *political system with freedom of investment* or some other related meaning, as Allott (this volume) recalls. He argues that a word like *democracy* may be used with little meaning, sometimes only to trigger positive judgements, and can be completed by 'whatever the speaker means by that'. We agree that this kind of strategy is among the core ones at work in manipulative discourse, since it triggers consent and adhesion not on the basis of proper arguments, but because of some fuzzy connotation. However, it is likely that the notion of *misuse of concept* needs further elaborations.

With the notion of "utility of interpretation" Chomsky (notably Chomsky 1989) actually addresses the misuse of concepts, which is claimed to favour the consent of the audience to thoughts or actions that are, in fact, incompatible with the concept actually encoded by the lexical expression, as with the word 'democracy'. In *Manufacturing Consent*, Herman and Chomsky (1988) detail their view on political manipulation; however, they do not cover the theoretical aspects in much depth.

Crucial imports are given by Manktelow and Over (1990), who correlate manipulation with problems of inferential processes, and Sperber (1982, 1985 and 1997), who addresses the problems of irrationality and covert communication from an informational viewpoint. Taillard (2000) and Blass (2002) develop in detail pragmatic aspects of manipulative communication and tackle the complex relationship between *informative* and *manipulative* intentions. In

particular, these authors argue that cognitive information processing is perturbed when dealing with words with vague, not understandable, vacuous meaning or whose content is inappropriate with regard to the context – a hypothesis that lies at the core of my analysis.

1.2. Manipulation as a type of language use

Every researcher approaching the field of manipulation, be it in linguistics, discourse analysis, psychology or political science, is aware of the vagueness, the semantic complexity and the lack of clear-cut definition for the concept *manipulation.* This is especially true when it applies to a specific behaviour towards other individuals, and to a type of linguistic and communicative behaviour in particular. Literally, to manipulate is to use one's hands to instrumentalize an object (*to operate or control by skilled use of the hands* according to the American Heritage Dictionary), and sometimes to change the object's original shape. Then to manipulate a human being may be about *using* a person, i.e. have that person adopt specific behaviours, to fulfill the needs and interests of the manipulator, regardless of the ones of the manipulated. But an individual, contrarily to an object, has a cognition that enables him to pursue his own interests; therefore, our first step is to admit that a manipulator first of all manipulates some aspects of human cognition, notably reasoning, checking for likeliness, emotions, etc.

The word *manipulation*, when applied to *manipulative discourse*, seems to be a kind of lexicalised metaphorical derivation. It is about using a device or strategy without which the speaker would not be able to change the addressee's beliefs and behaviour. Manipulation of an individual is directly related to *applying constraints*, in particular constraints that the individual is not aware of. These constraints act on the process of information treatment and are built-up with more or less efficient and sophisticated strategies – which remain of course *hidden* – and which aim is at misleading the hearer in a way or another. In effect, the commitment of the addressee to the propositions conveyed by the discourse must be sincere (or taken for sincere, for another type of manipulation, where physical and psychological constraint is used to guarantee a behaviour independently of the persuasive power, or lack of persuasive power of a discourse). The commitment of the addressee must be sincere while the propositions expressed are in fact problematic at several levels. Furthermore, hidden strategies are necessary because obvious and brutal force cannot gain *sincere* commitment.[2] Freedom of thought, or at least the illusion of it, is a necessary condition for manipulation.

The fact that manipulation lacks a clear-cut definition does not entail that the concept itself is completely unclear. To take an analogy from Wittgenstein, a country remains real and identifiable even though it can have fuzzy or controversial borders. Therefore, we will provide below an axiomatic working definition of *prototypical* manipulation, rather than explore the complexity of the intuitive notion attached to the word (we are talking about a phenomenon, not about the semantic content of the lexical item 'manipulation').

Even more complex is the notion of manipulative *discourse*. Concerning its definition, two options are available. First, *manipulative discourse* can be a *type of discourse*, therefore identifiable through formal features. Second, manipulative discourse can be a *type of language use*. These two lines of thought are not theoretically equivalent – but they may lead to close conclusions in the end.

If manipulative discourse is a discourse type, like *narration*, *theatre play* or *fairy tale* may be, then *either* some linguistic forms can be found *only* in manipulative discourses, or manipulative discourses are sustained by a unique type of structure, for example specific *argumentative* structures. In both cases, these particularities would provide a formal ground for manipulative discourse identification.

Yet it seems far more reasonable to assume that manipulative discourse is *not* a discourse type according to purely linguistic criteria. Manipulative discourses exist not because of formal features; they are produced in order for the speaker to achieve specific goals. Although some formal features are more present in manipulative discourses than in non-manipulative discourses, none are exclusive to manipulative discourses. The main criterion I will use is one of *intention* on the part of the speaker, an intention which is not cooperative in the Gricean sense (in particular regarding the respect of the maxim of quality). The speaker aims at giving manifestness to a certain number of assumptions to the hearer and have him consent to them, provided that they would be rejected under normal conditions. Manipulative discourse is therefore a pragmatic problem in my view. It is a type of *usage* of natural language, and can be identified only through notions like goals, intentions, and broader aspects of pragmatic processing, which, in turn, explain the quantitatively high presence of some formal features (some types of argument schemes and fallacies, some semantically loaded expressions, some connotative words etc.) because they are of some help in achieving the speaker's goal. Therefore one of the core problems of manipulation in language resides in the identification by the hearer of the manipulative intention through formal and non-formal features; when this detection fails, manipulation is effective.

This type of usage of natural language requires an extensive use of some types of argumentative devices, including formal and non-formal fallacies (see van Eemeren and Grootendorst 1992 for example). In short, I suggest that linguistic structures and formal phenomena taking place in manipulative discourse are provided by the speaker in order to trigger specific pragmatic processing. In other words, even if some linguistic elements are statistically frequent in manipulative discourse, these linguistic elements/structures aim at weighing on the pragmatic level, in order to affect the mental state of the hearer. Thus manipulation is not about using metaphors, or some particular syntactic structure, or some specific semantic feature of quantifiers, but about making them play a particular role at the pragmatic level.

Turning now to the problem of the propositional content of manipulative utterances, one notices that manipulation is often regarded as a sort of *lie* (for example, a recent account of lies in the public discourse of the USSR is to be found in Jaccard *et al.* 2003). The notion of *lie* is a complex notion, but seems to have a clearer definition, intuitively, than manipulative discourse. Although, there are cases showing that there is more in lying than simply uttering a proposition known as false or believed such. Saying '2 plus 2 equals 5' when knowing that it is false is a lie only when it is uttered in order to obtain some relevant advantage. If a kid says so simply to make a fool of another kid, and not for other purposes, he probably did not actually lie in the full meaning of the term. Regarding manipulation, as many scholars, I suppose that to manipulate implies to deceive in one way or another; although deceiving is slightly different from lying.

It may be argued that the manipulator sometimes actually believes in the proposition expressed. This question is very complex, and I address it briefly further down. In principle, I suggest that a manipulation in fact always entails a kind of content which is not fully adhered to by the speaker: not necessarily a lie proper, but something wrong on one level or another of the communication going on. For example, we know that beliefs can be entertained with various strengths; a weakly entertained belief stated with great authority can be manipulative since this situation implies a discrepancy between the (weak) strength with which the speaker actually believes in P and the (strong) strength in which he communicates believing in P. This is not a lie proper but is similar in nature. On the contrary, it can also be argued that a lie is not always manipulative. I will not address this question directly in this paper, but I notice that the notion of manipulation is not only associated with concepts like *lie* and *intention* but also the one of *interest* on the part of the hearer: a lie – or any defective statement – will be

manipulative if the hearer utters it in order to have the hearer adopt a behaviour consistent with the manipulator's interest and possibly inconsistent with the hearer's own interests.

Now, one may also argue that often, it seems that a true statement, sincerely entertained by the speaker, is communicated 'manipulatively' to the hearer. That might be true according to some weaker notion of manipulation, but I emphasize that a communication is manipulative when the speaker retains some relevant information, or provides the correct information but in order for the hearer to conclude that he should behave in a way which favours the speaker's interests, without being aware of it. Therefore, I consider that to manipulate is, firstly, to communicate the relevance of things that are not relevant by themselves, and / or retain actually relevant information.

A first way to study the propositional content of manipulative discourse is to adopt a moderate objectivist view. The main reason for this is the following: if A says to B the proposition P where P is true in reality but P is believed false by the speaker, it is hard to qualify A's attitude as simply *lying* (B certainly would not say, under any circumstances, that A was lying, although he was not communicating accordingly to his beliefs). Again, I refrain from drawing any strong conclusion about this.

2. Manipulation and truth-conditions

2.1. A definition

Within a hypothetical-deductive framework, I will now suggest a working definition for *manipulative discourse*. The advantage of this definition is that it overcomes the fuzziness of the intuitive notion through a light, and I think legitimate, reductionism. This basic definition of manipulative discourse has further implications that will be discussed below.

> A working definition: *A manipulative discourse is a discourse produced in order to persuade the addressee of a set of propositions P1…Pn of type T with appropriate strategies S.*

I shall address later the characteristics of the propositions conveyed by manipulative discourse (the 'type T'), positing that they are truth-conditionally or, better, truth-functionally defective (roughly, they are *wrong* in some way: false, unlikely, doubtful, inaccurate,

inconsistent with the common ground; and therefore should be rejected by the hearer under normal circumstances). The definition takes the following form, where G is the *goal* of the speaker *s*, λP stands for the *set of propositions conveyed*, CE(*h*) stands for the *cognitive environment of the hearer h* (the set of his beliefs)[3] and λS for the *set of strategies* aiming at achieving the belonging of λP to CE(*h*):

∃ G*s*: λP (where P = *truth-functionally defective*) ∈ CE(*h*);

G*s* → λS such as λS(λP) → λP ∈ CE(*h*).

There exists a goal G*s* on the part of the speaker *s*, which is fulfilled when a set λP of truth-conditionally / truth-functionally defective propositions belong to the cognitive environment CE of the hearer *h*. G*s* implies a set λS of strategies such as λS applied to λP entails λP's belonging to the cognitive environment CE of the hearer *h*.[4] Actually, the effect of the strategies are not granted: they are simply assumed by the speaker to work.[5]

Without the specification of the *properties* of the propositions conveyed and of the identification of the *strategies* used to convey them,[6] this definition would not be informative enough. Therefore, the properties attached to the propositions, as well as the communicative strategies involved, are wholly part of the definition I assume here. Let me begin with the properties of the propositions conveyed by a manipulative discourse, which are mostly about their truth-conditional or truth-functional value (I use *truth-functional* in a wide sense: implicatures are truth-functional since their truth-value is a function of the truth-values of the explicit content and of the contextual premises needed to deduce it).

I rely here on studies that have pointed out that totalitarian ideologies, and in particular ideological discourses within totalitarianisms, imply the commitment of the addressee to ideas that do not correspond with factual reality (Jaccard *et al.* 2003, but of course also Klemperer 1946/1975, Chomsky 1989 and many others). Since I stand upon a (moderately) objectivist framework, this question of *reference* and *truth-value* of the propositions is a very natural starting point for my study. This concerns first *propositions about states of affairs*, but also propositions about moral issues, which we may call *propositions about desirable states of affairs* (following Sperber and Wilson's Relevance theory).

2.2. Propositions about states of affairs

I assume that the main characteristics of a proposition P conveyed by a manipulative discourse, when P is about a state of affairs, is the discrepancy of P with its objective truth value or with the truth of intended truth-functional inferences normally drawn by the addressee when interpreting the current utterance (in particular *implicatures*, but also possibly *explicatures*).[7] More precisely, when a proposition P is conveyed by a manipulative discourse, either P is false (or half-true, which is equivalent to *false*),[8] or a relevant implicature I inferred from P and the context is false.

In Gricean terms, one might say that the speaker does not respect the maxim of quality, either on the side of what is said, or on the side of what is implicated. Within the framework of Relevance theory, I suggest that either P is presumed to be relevant, and therefore true, or at least assumed true by the speaker, or acceptable in principle, while it is not (which makes P a lie if the speaker is aware of that), or P is indeed true but contextual features lead the addressee to enrich the meaning of the utterance in deriving implicatures I – or in retrieving presuppositions – that are presumed to be true but which are in fact false (and known or suspected so by the speaker).[9]

This case, where P or I is false, is a prototypical case of manipulation. But the retrieval of some types of implicated information, presuppositions in particular, are less or not at all subject to cautious verification by the hearer. They are nonetheless derived truth-functionally since they are grounded on a truth-conditional sentence / utterance, even though they are not, or not always, part of the truth-conditional meaning of the utterance.

Again, it might be argued that the manipulators[10] are not always aware that the propositions they convey do not comply with reality or with the relevant scale of values. In other words, they may not be necessarily *lying*. For example when they had been persuaded with manipulative means, or when they are not aware of reality. But this would not fall into the scope of my definition: I suggest, then, that the speaker may be performing an action of persuasion but not a manipulation proper: he says what he actually believes in good faith, which implies that as far as he knows, there is no possible discrepancy between his own interests and the – freely evaluated – interests of the audience.

But the problem is that sincere beliefs may combine with insincere, fallacious reasoning. The ill-grounded conclusions thus constructed are then manipulative according to the above definition. More importantly, I speculate that the manipulator, in the strongest sense of the

word, is *always aware*, at least to some extent, of the falsehood of what he says or implies (or suggest: for example that P is relevant while it is not, regardless of its truth). This does not imply that the manipulator is fully or intrinsically insincere, since his awareness may be unconscious, if we see these terms as non-contradictory.[11]

Not even thinking of the cases where P is *known as false* but anyway *believed* by the speaker – a case of psychological trouble –, the view according to which a manipulative discourse is based on the violation of the Gricean maxim of quality is not satisfactory: it is far too rough to consider that *to manipulate* is simply *to say what you don't believe* or *to say something you don't have good reasons to believe*. There is a complexity that lies behind the production of a manipulative discourse that cannot be captured through Gricean maxims. For instance, as said before, communicating that P is relevant while it is not is not a proper violation of the maxim of quality.

There is a difference between what the speaker actually believes and what he may *wish* to believe: manipulators are certainly involved in self-persuasion as much as they are involved in persuasion of the hearer. A manipulator may find it more desirable to be entertaining beliefs that promote his ego or positive self-image, or power, rather than having thoughts and behaviours in accordance with facts. No doubt in some (many?) cases of manipulation, the manipulator is more concerned with being admired and empowered rather than with the need for correspondence of his beliefs with reality.

This question is far too complex for me to address here; it is about the details of the production of a manipulative discourse, and about the psychology of the manipulator, while I focus on the mechanisms of detection of deception and how they are defeated by means of manipulative strategies. However I suggest that contrary to psychotics, manipulators (in my restraint sense) are always aware, to some degree, of the falsity or doubtfulness of the propositions they communicate, as I said before. This is why manipulators often omit some necessary parameters or premises in their arguments and promote censorship and repression against intellectuals.

What is crucial is the fact that the *goal* of manipulators, whatever they themselves believe and regardless of their particular psychological condition, is to convince the addressee of a proposition or set of propositions that should be ruled out by normal information processing and reality checking. The manipulator short-circuits the process of reality checking and disturbs normal information processing. Manipulators are *more* than liars, since they aim,

among other things, at producing dogmas that will in turn provide an axiomatic set of beliefs in the reasoning material of the manipulated.

2.3. Non-factual propositions in manipulative discourses

Quite often we find manipulative utterances that are not about states of affairs but about moral issues and principles. Such utterances are not confronted with the addresses' conceptual representations but with the addressees' moral values and culturally-sensitive representations. In other words, these propositions deal with representations of *desirable states of affairs* and not with *actual states of affairs*. The consensual way of discussing this kind of statement is to talk about *social* or *cultural values*, which are themselves grounded on deeper moral values.

In fact, with such propositions the mechanisms of manipulation are the same as with truth-conditional propositions except that a moral proposition P is not evaluated regarding its *truth* but its *acceptability* to the ethical values and cultural background of the target audience. Moral statements and propositions about desirable states of affairs are the easiest type of propositions conveyed within a manipulative discourse, since reality can be ultimately checked, whereas moral values are not stable in any comparable way.

These points lead to another hypothesis: totalitarian ideological manipulative discourse is most efficient in societies where some conditions are met, especially a state of crisis. For example, economic recession, war, and post-war situations are factors that favour less stable moral judgement for individuals and open the way for a change in moral values. The fact is that when a society or system has not been able to provide appropriate quality of life and the possibility of human relations, individuals are likely to conclude that the values that ground that particular society or system are not appropriate. This plants the seed of doubt and the will to change societal conditions. Needless to say, such change is not trivial.

If socio-cultural values, like the commitment to democracy, equality, and rights, are weak, other contradictory propositions will bring about a change in the addressees' cognitive environment that will allow the development of new beliefs. Remember that assumptions always have a certain degree of certainty, a certain *strength*. This applies to moral statements as well as to factual statements. For example, a belief like *killing is bad* is generally stronger, at least within occidental cultures, than a belief like *stealing an apple is bad*. This is a simple way of talking about scales of values.

When the proposition has a moral meaning or has moral implications, it is evaluated with regard to the prevailing ethical beliefs. Within a public discourse, the speaker takes this into account as the confrontation of the statement with the prevailing values in the culture where the ideology intends to promote itself and develop. We can define the moral culture as follows:

The moral culture C is a set of assumptions about desirable states of affairs prevailing in a given population.

We note that C *sanctions the acceptability of moral propositions* conveyed within a public discourse.

As I said, if C is weakened, for example in the case of social crisis, new moral statements can be entertained more easily. In other words, *the weaker C, the stronger P (where P is manipulative and where P is not in accordance with C), the more successfully P is accepted.*

However, manipulators can also exploit some moral assumptions in the same way as they can exploit some aspects of obvious reality. Just as they can extract and present half-truths, they can extract and exploit half-moral 'truths', or only part of a moral statement, and distort it. For example, by exchanging different moral assumptions within the prevalent scale of value, they may receive guidance for otherwise immoral actions, as when assuming that a superior good justifies prevailing over a supposed less important good. This was the case when the killing of disabled persons by the Nazis was intended to achieve the alleged superior good of a pure and healthy society or when the killing of innoncents by the Tsheka was done in order to build a society without certain social classes. All these aspects of manipulation are well documented by other scholars (in particular in the philological study of Klemperer 1946 / 1975, as far as Germany is concerned) and I will not comment them here.

2.4. Three kinds of proposition-evaluation

A *contextual evaluation of the propositions* is rendered by the hearer in order to guarantee that he can reasonably consent to the propositions provided. *Either* we evaluate whether it is reasonable to assume the correspondence of the supposedly intended meaning with what we assume about reality, *or* we evaluate whether the moral consequences of the intended meaning are compatible with non truth-conditional beliefs.

The first type of evaluation concerns the consistency of the proposition with background assumptions about reality. The second type of evaluation concerns the consistency of the proposition with assumptions about desirable states of affairs – moral assumptions. The two types of evaluation can both occur through the process of interpreting an utterance, when the interpretation results in a combination of truth-conditional propositions (about actual states of affairs) and implicatures that carry moral significance (about desirable states of affairs).

The same applies to sets of propositions together forming fallacious arguments. When fallacies (formal and non-formal) are provided in good faith, one cannot talk about manipulation. But when arguments are deceptive (fallacious) and are used to promote other defective propositions, there are good grounds to *suspect* a manipulation. The *bandwagon* fallacy and arguments of authority, for instance, are particularly present in ideological manipulation.[12]

Our cognitive hypothesis is that the manipulator acts *on the very process of this evaluation*, leading the hearer to block his own natural process. There are therefore specific strategies that the manipulator exploits to obtain such consent (which I address below).

One final issue concerns the fact that manipulative utterances and propositions[13] are often *vague*, *excessively metaphorical*, *pseudo-mystical* and *confusing*. It is a well-documented fact that manipulative discourses often use these kinds of linguistic devices in order to call for irrational consent (see for instance slogans in the Stalinist era, odes to charismatic leaders or, more simply, the fight against 'rationalism' under the Nazi regime). However, these aspects are not part of the *necessary* characteristics of the propositions conveyed by a manipulative discourse, since, on the one hand, the linguistic form is relatively incidental and, on the other hand, many vague statements are not at all manipulative. It is worth repeating that, in my view, manipulative discourse is not a discourse type that could be identified through observation of linguistic features alone. Instead I suggest an hypothesis regarding the *pragmatic effects* of fuzziness and metaphorical excess, among them the generation of a trouble that – paradoxically – may favour the confidence towards the speaker in certain conditions.

3. On manipulative strategies

The strategies used by the manipulator in order to block *truth*, *likeliness* and *acceptability* checking, as well as *consistency* checking, are first linked to the necessity of blocking the identification of the *manipulative intention* of the speaker. Manipulation is a case of covertly transmitted information – it would be self-defeating if the addressee recognises the speaker's intention and came to a certain conclusion. But manipulation is even more: it is about blocking one's rational device so that certain operations about beliefs are accomplished by the manipulator himself, I suggest.

We consider that the strategies used by the manipulator are of two general kinds: *local strategies* and *global strategies*. Local strategies are those used to constrain the interpretation at the level of utterance processing. Global strategies are those used to create adequate social and psychological conditions to obtain irrational consent. Global and local strategies can be both linguistic and non-linguistic.

3.1. Local strategies

Linguistic local strategies are already well-known, in particular because of the work of German philologists during or shortly after WWII.[14] Most of the linguistic local strategies used to defeat the natural rules of utterance processing are oriented towards the production of *fuzziness* in general: a fuzzy thing or thought, presented with great arrogance and authority, creates a problematic double-bind. I come later to this problem of double-bind and pseudo-resolution of the double-binding, which, I suggest, lie at the core of the manipulative process. Other strategies involve presuppositional assertion, misuse of concepts, and pseudo-mystical discourse.[15]

Rhetorical devices such as rhetorical questions are quite efficient, since such questions carry strong presuppositions. A question like (1) carries the presupposition that anyone disagreeing with the speaker is a traitor:

(1)   Which traitor would give our homeland to the imperialists?

Some questions such as 'what else can we think?' presupposes that the answer is obvious: 'Nothing'. The implicature that this 'nothing' means 'nothing at all' and not 'nothing that I know about' is automatic and is due to general principles of communication (Gricean maxims or principle of relevance, Sperber and Wilson 1995).

'Misuse of concepts' is a problem already mentioned, and is addressed in more detail by Nicholas Allott in this volume. An interesting example can be found in Le Pen's speech when he won the first round of the French elections in spring 2002. He talks about miners and workers of the steel industry ("*métallos*", a word which connotes hard work and precarious living) in a sentence that could translate as 'Don't be afraid, you the humble, the excluded, the miners, the *métallos*.' Talking about *miners* raises a referential problem since most coal mines in France have been closed for years, and nearly all the miners are now unemployed or have converted to other activities. Therefore, the miners form a very little part of Le Pen's audience when he speaks to the Nation. But he addresses the audience itself with these words "you the humble, the excluded, the miners". Le Pen is talking about virtually non-existing referents and asserts their existence by presuppositional means. The word *métallos* triggers a related problem: steel in France is no longer a leading industry and only a few of the former *métallos* are still working as such nowadays. In the interpretation of such utterances, the hearer gets something like 'you, the *miners*, whatever this means, you the *métallos*, whatever this means', while only connotative presuppositions or implicatures are kept and, one should say, *driven* by the words immediately preceding: "you the *humble*, the *excluded*". Many other features of this utterance deserve interest, among which the fact that Le Pen, so to say, presents himself as the one who talks to the "excluded", which triggers positive and empathic feelings also in the mind of the non-excluded people.

Another well-known misuse of concept was made by the use of the words *hero* and *heroism* in both the Third Reich and the Communist countries. There, the hero is the person who will follow the rule of the dictator or the party and act as requested by the power in place. Killing of innocent and unarmed people has been many times called *heroic*. The same applies to the German *Kampf*: the 'noble fight'. Every activity of everyday life was a *Kampf*, a noble fight for a noble cause. These facts are neatly addressed by Klemperer (1946 / 1975).

In the USSR, it was common to use *bandit* (approximately the same as the English and the French *bandit*) to designate not only a robber, a criminal or a gangster, but also a soldier of any anti-communist guerrilla movement (this use of *bandit* was first applied to Chang-Kai-Sheck's movement in China). The word *bandit* has even extended to describe any anti-communist activist (as Shmelev 2002 points out). The word *comrade* ("*tovaricht'ch"*) is also a good example of the misuse of concepts.

In addition, religious-like discourse plays a very important part in ideological totalitarian manipulation. It implies that the speaker is a God-like being, a Saviour-like being. The very

beginning of Le Pen's address to the nation is a good example: "Enter in hope, be confident." The speaker asks for faith instead of rational agreement (it must be added that Le Pen uses the French word *espérance* ('hope'), which is used in religious contexts, instead of the usual word *espoir*).

Unmotivated analogies, metaphors, and vague terms, together with the effects of global fuzziness, make it complicated to decode the real information that lies behind the discourse and the articulation between these pieces of information. Thus the hearer is made less able, or even unable, to identify logical fallacies, half-truths, and other defects of the discourse. Fuzziness also plays a role in feelings of incompetence in the hearer's mind, which is very important in the logic of manipulation. In all cases, the addressee is put in a state of confusion, because he can't trust his own intuition that the discourse is meaningless. What he believes instead is that he is not competent enough to understand it.

Non-linguistic local strategies *accompany* the production of the discourse. They are elements that contribute to produce a situation where the hearer is under the lead of the manipulator. Such strategies are numerous, from the general attitude of the speaker to prosodic features and intonations,[16] and to the triggering of emotional reactions. When the discourse is written, it is accompanied by pragmatic effects coming from the typeface, the organisation of text on the page, images, etc. Wilke (1998) makes it clear that these constraints apply to journals and newspaper layouts.

3.2. Global strategies

Well-known linguistic global strategies (manipulative strategies that rely on language use but which are not directly provided in some particular manipulative discourse) include:

- Spreading and repetition of specific connotative words (words that trigger under normal circumstances implicatures (or presuppositions) with symbolic weight);

- Generalisation of a new terminology;

- Elimination of some lexical items from public discourse;

- Unmotivated or misleading analogies;

- Acronyms, abbreviations, numbers;

- Naming of elements of the everyday environment

Specific connotative words include military terminology, which presupposes (or triggers the implicature) that the social individual is, metaphorically, a soldier who does not question orders and finds happiness in a state of instrumental submission. New words or partially new constructions, such as the well-known generalisation of the distance prefix *ent-* in German, used in order to build new verbs and substantives, are linguistic global strategies of manipulation.[17] When stabilized in the public lexicon, these expressions contribute to legitimate as 'normal' and 'conventional' a range of concepts – in particular (but not exclusively) regarding actions that the target audience is expected to achieve – which would be seen as strange or exceptional under other circumstances. When this point is reached, one can hypothesize that the individuals do not really perceive any longer that these namings were at first imposed by the power, since any lexicalized item falls in the cultural common ground shared by the community as if *sui generis* (the perception we have of words and meaning are not diachronic but synchronic: the history of their implementation in the idiom is not salient at all when we use language automatically).

Unmotivated analogies which are constructed with discourse but materialised by institutions may also appear as linguistic global strategies of manipulation. For instance, one may question the classical links established between health, sport, youth and the army, when there is an obvious lack of natural motivation in the reference to the military activity. This is especially true in twentieth-century armies, where no swords, spears and halberds are used. There was, under the Nazi rule, a hybrid ministry dealing with sport and the army; in the same vein, students were forced to practice something called *Wehrsport* ('military sport') on Tuesdays afternoons, and there was even a brand of cigarettes called *Wehrsport* (again a fact reported by Klemperer 1946 / 1975). These kind of things create a context where assumptions of the type *the army is a good and healthy sport* are present in the hearer's cognitive environment, even if they are often unconscious, unreflective beliefs.[18]

The creation of acronyms and abbreviations of all kinds, the loss of personal and institutional identities through the use of identification numbers, as in the names of schools and other institutions in communist countries (like 'school 22' or 'hospital 17') all contribute to create a global atmosphere where depersonalisation is progressively integrated and accepted as conventional and normal. The Soviet naming of streets, towns and villages like 'progress', 'steel plant', 'electrical power' and others, affects the everyday environment of the manipulated and creates a positive image of the human being as a worker. In addition, it implies that human existence is not only dedicated to a function in the system, in particular a

working function, but also *defined* by it. This form of global alienation results from the depersonalisation of the individual through naming.

There are, I suggest, two main devices of global non-linguistic strategies.

The first one is a purely social device: group pressure (the individual tends to comply with the most salient opinion within the group he belongs to) – a very well-known topic within social psychology that I will not address here.

The other device seems to me more crucial as far as consent to defective propositions and arguments is concerned. It is the achievement, in the hearer's mind (and in public opinion), of a particular *image* by the manipulators, who want to make themselves appear super-competent, either because they manifest in one way or another that they have some information that the hearer does not have, or because they are viewed as uncommonly skilled. In prototypical cases, for example in a typical twentieth century dictatorship, the cult of personality focused on dictators is extreme. The over-competent image of the manipulator has a non-trivial consequence: the dictator appears as a superior mind, a universal genius, a being who knows more than any ordinary human about the world, destiny and society, which entitles him to establish transcendent-like dogmas. The manipulator in totalitarian societies can be compared with a God-like figure (to whom an alienated form of worship is dedicated) The same arises when religious fundamentalists preachers want to be taken for prophets, which can link to suicide bombings or whatever abnormal behaviour.

This image is built up in order to achieve a specific goal: the hearer must be *confident* in the leader, this confidence being a condition of obtaining sincere consent, becauses it creates the belief that the manipulator acts in favour of the individual's own interests.

In democratic societies, where manipulation is likely to happen as well, other schemes may trigger confidence for the leader. The reason is the one which lies behind any persuasion process, and we take it axiomatically: if hearers are confident in speakers, they tend to believe them. The simple reason for this is that confidence lowers critical thinking and is, then, is energy-sparing (I assume, following Relevance theory and others, that the human mind is geared at following paths of least-effort). Defeasing a strong assumption about the speaker deserving confidence implies a cost.

As confidence is a matter of scale, I will assume the following:

The more confident the hearer is, the less critically he thinks, and the more efficiently the manipulator is likely to achieve his persuasive goal.

Very likely, the human mind has a 'source-tagging device' evaluating the reliability of the source of the linguistic communication (reflected in various modal and evidential expressions grammaticised or lexicalised in natural languages); a prominent source is 'credible authority'. It follows that manipulators have to establish credibility, and it is common sense that the building of an over-competent image is a key to this process of gaining confidence (see in particular Sperber 2000).

The important fact, making manipulation fundamentally different from rational persuasive effects, is that *the grounds for which the hearer is confident are not rational*. It is a matter of fact that, generally speaking, children tend to believe their parents, non-specialists *tend* to believe specialists, patients tend to believe doctors, and so on. In these cases, the grounds for confidence are rational.[19] There are good grounds to think that the speaker knows more about the topic of speech that the hearer, and this leads naturally to confidence 'by default'.

But there is another condition in which confidence arises. The hearer must have no reason to think that the speaker is deceptive or that the speaker doesn't wish the hearer's good. In the examples mentioned (the doctor, the parent), it is generally the case that the speaker is benevolent towards the hearer.

The condition of benevolence can be exploited by the speaker. When benevolence is perceived as *extreme*, which means that the speaker is perceived as pursuing the interests of the audience *more* than he's own interests, affective reflexes arise in the audience. Again, the most extreme case reveals how the speaker gains the confidence and even a form of (sometimes very strong) affection from the individuals in the audience.

We normally trust more easily people we love and who claim to love us, unless we have good reasons to think that they are not competent. If A believes that B loves him, A assumes that B wishes A's good, and that B's actions will be oriented towards promoting A's interests. If A is a group and B a dictator exhibiting his love and concern towards his people, then it is more likely that A will believe B. As an example, Perón, in a famous speech, said he was renouncing the rank of general and personal glory for the sake of the people, an argument certainly very efficient in obtaining confidence.[20] This happened also when Pétain asked the people to have faith in him (French capitulation speech), when Stalin was called 'Father of the peoples,' when Ho Chi Minh was called 'Uncle Ho', and so on. Thus, appeals to emotion and pseudo-mystical discourse gain confidence through emotion-sharing, in turn communicating

destiny-sharing, from which the hearer expects that the speaker will fulfil his existential needs.

If we also consider that A believes that B is super-competent, then all the ingredients for successful manipulation are in place, provided that B favours his own interests and not those of the audience, and that he is ready to mislead the audience for this. This approach can be applied to manipulation within sectarian cults as well.

It may be tempting to count among global non-linguistic strategies physical force and use of power, inasmuch as they strengthen group pressure and sustain an ideology where enemies are identified and subsequently punished or destroyed. This matter is complex, but I would rather suggest that force gains sincere consent only under very specific conditions and, therefore, is manipulative only when generalized repression is active. In effect, generalized repression may lead to a psychological economy of consent where an individual finds it less costly to adopt, step-by-step, the thoughts expressed by the authority, rather than trying to resist (Stockholm syndrome is an extreme case of this mechanism, although all but rare, where the dominated comes to love the dominator regardless of the harm caused by the dominator).

But when force and power become the most important or only tool of power preservation in a community, another dynamics is put in place (see Ilie in this volume). Relying on force and repression, the power in place does not need that much to convince the audience. Among leaders and members of their staff, fewer people sincerely believe in the ideology while becoming more and more aware that repression and force are the main guarantees of their privileges. These situations, where only a small minority of people believe in the public discourse (maybe in some cases almost no one) or at least fake to do so, while the rest remain in submission because of force (as was the case in the last decades of Ceaucescu's reign in Rumania), generally signal the close collapse of totalitarian regimes. Khrushchev's de-Stalinization was perhaps intended to avoid this danger in post-Stalin Russia.

4. The central mechanisms of manipulation: trouble and resolution

Among all these strategies, linguistic ones are crucial for the manipulator to obtain sincere consent. And of all linguistic strategies, we suggest now that *fuzziness*, in a wide sense, plays an important role in the process of manipulating. Let us elaborate on this point.

First of all, fuzziness is not only about vague terms, complicated sentences and overuse of metaphors. One of the most important strategies used within manipulative discourses is the creation of a *global* fuzziness, where even parts of the discourse that seem clear and simple lead to interpretative problems for the addressee. Simplistic, and usually inductive arguments violating logical rules, and which create unmotivated generalisations, can be considered as contributing to the global fuzziness or confusion of the discourse,[21] despite the fact that they are intended to imply a particularly 'clear' and wise thinking.

Simplification creates another problem. When facing this kind of discourse, the addressee is troubled, since his cognitive ability to clearly understand the speaker is disabled in one way or another. In such cases, the hearer may well be aware of an argumentative problem, and we think this is generally the case. However one must explain why the manipulated comes to adopt a simplistic view on complex problems.[22]

The manipulative discourse, at least within totalitarian regimes, is *impressive.* First, it creates non-propositional effects on the addressee, such as emotions that give rise to the assumption, on the addressee's side, that the speaker is *sincere* and that he shares the worries of the addressee. Second, the discourse is impressive because the speaker makes use of concepts and arguments that the target addressee would not use easily, or because it draws upon reasoning that complies with some personal intuition or opinion on the part of the hearer, thus providing an appearance of rational justification for these intuitions or opinions.

In many other cases, the trouble does not arise from simplification (and fallacies) but from obscure, vague, metaphorical, mystical expressions. The addressee faces a conundrum: one the one hand, his natural ability to understand clearly is defeated (he has an assumption about his self-incompetence), and, on the other hand, he is keen to believe in the speaker's word, because of these non-propositional (emotional) effects, combined with the constant request, on the speaker's side, to have faith in him. In addition, there are social devices implied in the communication, such as group pressure, which can lead the hearer to overcome a first impression about the problematic properties of the utterances. This phenomenon is likely to be related to the kind of situations many of us have once faced, be it during childhood or not, when we happened to laugh at a joke we didn't understand just because it would be face-threatening not to or because we wouldn't like to risk a self-exclusion.

Moreover, the addressee is led to believe the speaker's word, since it is communicated to *solve* the double-bind, which is thought to be due to the hearer's incompetence but which is in

fact created by the manipulator himself. The hearer abandons some of his cognitive abilities for the sake of higher values presented by a discourse originating from a seemingly higher intellect. This way, the addressee is in a position of moral, intellectual and psychological dependence towards the speaker, who in turn appears as a saviour, a genius, a 'God-like' being.

One more word should be said about emotions. Quite often, the discourse is accompanied by other emotional devices than prosody and intonation. Emotions are also triggered by devices at the level of the general attitude of the speaker, and relative to the propositional contents of the utterances, as when the manipulator speaks about concepts that call for fear and hope. The addressee perceives the speaker's emotion, fake or real, which seems fair and motivated, since the addressee belongs to a group assumed to be integrated in a common fate with the speaker, and the emotion either comes to be shared by the addressee (for example fear) or a related emotion is triggered (for example when *enthusiasm* triggers *hope*).

5. Conclusive remarks: The manipulative intention and the theory of mind

One of the crucial problems of manipulation is the identification by the manipulated of the manipulative intention of the manipulator. How is it possible that normal humans, equipped with rational devices, stop using them correctly when confronted by an efficient manipulative discourse? Each of the aspects that we have discussed, plus a number of others that have been described by researchers, play an important role. A few points need to be explained in order to address the question of *intentions*. Relevance theory admits two levels of intention on the part of the speaker, which must be acknowledged by the hearer in order for the interpretive process to take place.

First, the relevance-theoretic idea of *ostensive-inferential communication* involves the need for the identification by the hearer of the speaker's *communicative intention*; the speaker makes it clear to the hearer that the stimulus is *intended* to communicate something to him. The recognition by the hearer of the speaker's communicative intention triggers *expectations of relevance* (the expectation that the utterance is adequately formatted to communicate relevant assumptions at a lower interpretative cost), and therefore allows for inferential processing to begin.

Second, the communicated material is not simply decoded but is pragmatically enriched (following a deductive non-demonstrative scheme of information processing). The full interpretation is a set of assumptions corresponding to what the speaker 'means' by the utterance. This set of assumptions represent the *informative intention* of the speaker, and includes any relevant proposition derived during the interpretative process. In order to deal with other types of intentions, not specifically communication-oriented, we would need a rather complex psychological apparatus. A few tools derived from Relevance theory and cognitive science regarding covert communication and detection of deception may well help (see Dennett 1989 and Sperber 1997 in particular).

We notice that there is no reason for which the deductive inferential process should stop when an interpretation is found for an utterance; that is, when the informative intention of the speaker is retrieved. We suggest on the contrary that the inferential process goes further, in order to find out about other types of information, notably the goals of the speaker and related intentions of the speaker, plus their consequences in interaction. This happens, we think, when the hearer has reasons to suspect that the communication which is going on may lead to relevant consequences. Relevant consequences may be ones that will help to improve the hearer's knowledge of reality (this happens when the interpretation of an utterance leads to further conclusions when confronted with other elements of the hearer's cognitive environment), or that will help escape from a risky situation in the interaction itself. This happens when the hearer suspects the speaker is playing a role, fake, or is being non-cooperative in some way, which allows for a hypothesis about the speaker's having hidden and potentially harmful intentions. This, and any weird utterance in context, or any utterance having weird consequences, or a proposition that seems to undermine assumptions we hold, all need to be evaluated in terms of *risk* for the hearer *or* improvement of the hearer's view of the world before the information comes to be held as a belief or on the contrary rejected.

Dennet's "intentional stance" is based on a similar view, although with other concepts, and, seemingly, under the assumption that the process is fundamentally reflective:

> Here is how it works: first you decide to treat the object whose behaviour is to be predicted as a rational agent; then you figure out what beliefs that agent ought to have, given its place in the world and its purpose. Then you figure out what desires it ought to have, on the same considerations, and finally you predict that this rational agent will act to further its goals in the light of its beliefs. A little practical reasoning from the chosen

set of beliefs and desires will in most instances yield a decision about what the agent ought to do; that is what you predict the agent *will* do. (Dennett 1989:17)

The core idea is that the individual acts in accordance with his goals and intentions. Therefore, actions are the key to finding out about intentions. Relevance theory has an even more radical view on this, assuming that a specific device in the mind is dedicated to the detection of intentions: the *mindreading module* (after the well-known works of Baron Cohen). The role of such a device is the same as the role of rationality in Dennett's view, and is also the same as any variant of the *theory of mind*, itself a variant of what used to be called *popular psychology*. All these trends admit in one way or another that the human mind has some sort of 'reflexive knowledge' or capacity to elaborate about itself, positing similar mechanisms to other individuals.

The question we need to address is *why is the manipulative intention not detected* by a critical amount of hearers although there are obvious or good reasons to suspect manipulation? This can only be answered through a more fundamental question: does human cognition work normally or not when interpreting a manipulative discourse? These are problems that lead beyond linguistic issues, but are directly related to the question of the interpretive processing of utterances (they are pragmatic questions). We know that these aspects of things need further elaboration and research from other disciplines, notably standard psychology. But there are some things that can be said about that from a cognitive viewpoint.

First, one may hold that normal cognitive abilities are simply *exploited* by the manipulator. It is in effect certain that presumption of relevance (presumption that the utterance is relevant and thus deserves interpretation), or any cooperative principle, is in a way exploited by the speaker – the communicative act presents itself as a *normal* one. This is certainly true to the extent that the manipulator communicates the presumption of relevance of his utterances (or the presumption that he is cooperative). However, it is far more crucial to remark, first, that the hearer evaluates the possible relevance of the utterance, and can well decide not to interpret it fully (that happens all the time when someone talks to you and you think that his discourse is irrelevant, a phenomenon quite common when a scientist listens to a conference paper by a scholar belonging to an opposing framework, or when a French deputy drowses when listening to the Prime Minister during debates at the Assemblée Nationale). Second, and more importantly, the hearer can actually presume that the utterance is indeed relevant– assuming then that the speaker is both sincere and rational when communicating the

presumption of relevance carried by any utterance – but nevertheless finds out that the utterance is in fact irrelevant. In typical cases of manipulation by discourse, the fact is that these checkings that normally take place during utterance interpretation do not take place the way they should.

The first deception is at the level of what Relevance theory calls the *communicative intention* and not at the level of the information communicated itself. In fact, what is exploited is not simply the cooperative principle or the presumption of relevance, it's the *false belief*, entertained by the hearer, that the speaker is benevolent, cooperative or relevant.

But the interesting point is that when the addressee encounters a discourse with falsities and fuzziness, he does not (always) consider the discourse irrelevant. A mix of badly articulated fuzzy propositions, to take an extreme case, should not create consent, precisely because the presumption of relevance is not satisfied. How come the addressee questions his own ability to find proper and rational relevance to the utterances he processes?

Many devices that we need to discuss, such as the god-like image of the manipulator, merely open the way to a more global hypothesis on the topic of manipulation and cognition. In short, we suggest that the way to avoid identification of the manipulative intention resides in the god-like or super-competent image of the speaker. And this has a direct link with intention-recovery.

But there may be a more direct way of explaining the success, when it happens, of manipulation in discourse. Given the fact that the hearer accepts the presumption of relevance, it becomes more costly to reject the propositions expressed after that first 'agreement' to process the information. There comes the problem of cognitive dissonance, exactly as exploited by telephone marketing trainers when they tell employees to ask first a question the potential customer is likely to answer by *yes* (for example "Do you read newspapers?" or "Do you have windows in your home?"). The intuitive awareness of this phenomenon may also explain the fact that sometimes, an individual will prefer *not to be exposed to some speech* as a self-protection reflex, in order *not to be persuaded*, just because it's complicated in the end to resist continuous persuasive speech, in particular when lacking the intellectual tools of critical evaluation (notably because of poor or inadequate education).

Whether this mindreading ability is grounded on the automatic and non-reflective discovery of intentions or on a (partly) reflective theory of mind is not important here. The interesting point is that any theory of mind or intuitive heuristics for the discovery of

intentions applies first to *ordinary humans*. But attributing intentions, and in particular very complex sets of intentions, to a God, or non-ordinary, super-competent, human being, whatever this may exactly involve, becomes problematic under the assumption that the ability of intention recovery is simply not applicable to the domain. Of course, recovery of intentions is an anthropomorphic ability that applies to other species and even to objects. But regarding a being that is felt 'higher' than an ordinary human, a god, or a genius, the question is not that simple. If the hearer takes for granted that the manipulator is benevolent and more competent than him, he may simply not consider himself able to attribute precise intentions that would not fit this scheme. In particular, a negative intention such as a manipulative one may not be detectable if the hearer believes strongly in both benevolence and super-competence. Moreover, intentions attributed to a super-competent being are likely to be weakly attributed, precisely because the hearer assumes a difference of status: while the speaker is super-competent, the hearer, by contrast, is under-competent. The statement that God's intentions are not understandable, as folk wisdom claims, reflects this aspect of things, and applies to any dictator believed to be benevolent and super-competent. In a crucial way, the manipulator achieves his goals when he makes the audience believe this. The result is a form of faith on the part of the addressee.

There is however an even more fundamental question: how come a given individual accepts the idea that he should stop evaluating critically the propositions made by the speaker? There is a rational factor: if the speaker shows super-competence then it is reasonable to accept his views. But under normal circumstances, we know that super-competence – or simply competence in a domain where the speaker is not competent – needs to be constantly demonstrated by the speaker, otherwise confidence is lost (think of the doctor who suddenly starts saying weird things). We can add a number of psychological and sociological factors such as group pressure, again. But the clue to all this is that the speaker builds an environment for the hearer which makes it *more efficient* for the hearer to adopt the views of the manipulator. Resisting group pressure, evaluating complex propositions when plunged into interpretive trouble, etc., is cognitively costly. Therefore, the fundamental assumption that I would hold on the success of manipulative discourse is the following:

The more costly it is for the hearer to retrieve correctly the information communicated, and to evaluate the truth, the likeliness or the ethical acceptability of it, the less likely the hearer is to resist manipulation.

It is pretty economical to rely on the other's opinion provided that he's thought to have the right opinion, whatever this opinion may be.

Of course, not everybody is manipulation-prone. It *needs* to be accepted, *sincerely accepted*, that the manipulator is super-competent, and that, therefore, an ordinary human has no authority to engage in normal cognitive processing of the manipulator's discourse. Many people are not totally manipulation-prone. However, when the group pressure is extreme, notably because of fear, any human is oriented towards developing a discourse that complies with the manipulator's discourse. When this lasts, many people tend to adopt, by necessity of survival, the basics of the official dogmas for reasons of economical and safer ordinary life.

Religion itself, when not used as a fundamentalist and exclusive ideology, can generally be counted among efficient counter-powers to manipulation, since the psychological 'throne' of an exclusive God is already occupied.[23] This is one of the reasons why religion often vigorously opposes totalitarian regimes and, in general, is a central concern for dictators (an obvious fact in communist countries and in Nazi Germany, where the Churches were a target, in particular the Roman Catholic Church, despite the ambivalent attitude of the Vatican at that time). Sometimes, however, dictators try to get the support of religion for strategic goals, since generally speaking a religion can give moral caution to his actions, and therefore makes them easily acceptable to people. For example, Pétain was anxious to get the backing of the Vatican, and the contemporary right-wing extremist Le Pen has its annual party meeting on Joan-of-Arc day, which is intended to promote an alliance of right-wing extremism and Catholicism. Juan Perón, in Argentina, got the support of the Catholic Church, which was crucial for him, but once in power, his connivance with the Church was no longer important and he started to promulgate reforms which were against Catholic views (notably on divorce and prostitution).

And of course, people are much less prone to manipulative discourse when they are aware of some of the central mechanisms involved.

It still needs to be said that what the French call *langue de bois* ('Wooden language': a string of ready-made sentences, generally ideological, without any substantial meaning, typical of communist regimes in the last decades of their lives) is not efficient as a manipulative device, since emotional support is not present. When wooden language appears as it really is, a discourse without significant content, the addressee is no longer manipulated. This is why wooden language generally appears only: i) when police and brutal pressure

enforces a behaviour of submission; or ii) when no other political force is present to challenge the power in place. In both cases, the discourses no longer need to be convincing, they fulfil other needs.

For a totalitarian regime to emerge, it is unavoidable that a significant part of the audience actually believes and supports its promoters. Active manipulation may decrease in intensity when the regime is actually supported by repression against opponents. Some claim that when this happens, the totalitarian regime is entering into a self-destructive phase. Although apparently clear cases of this phenomenon can be identified, for instance the fall of the Apartheid regime, it would probably be too strong to posit a generalisation regarding this point.

As a conclusion, I suggest that manipulation in discourse is clearly a *pragmatic* and *contextual* problem, where the notion of context is understood as the subset of the hearer's cognitive environment which allows for the interpretation to be constructed. As such, manipulation involves cognitive processes; my main hypothesis is that normal interpretive processes are troubled at the level of *intention recognition*, which involves a specific cognitive device (mindreading module). I suggest that this trouble is achieved with a set of converging strategies leading the hearer to problems of understanding, – notably, but not only, because of unclear propositions and arguments. I also note that many aspects that were evoked in this paper require much further elaboration and research.

References

Allott, N. and Rubio Fernandez, P. 2002. "This paper fills a much-needed gap". In P. Afuta, A. El Ghali and F. Toussenel (eds), *Actes de l'Atelier des doctorants en linguistique 2002* (Université Paris 7), 97-102 (unnumbered pages).

Allott, N. 2002. "Relevance and Rationality". *UCL Working papers in linguistics* 14: 69-82.

Berning, C. 1964. *Vom Abstammungsnachweis zum Zuchtwort: Vokabular des Nationalsozialismus*. Belin: De Gruyter.

Betz, W. 1955. "The National-Socialist Vocabulary". In M. Baumont, J.Fried and E. Vermeil (eds), *The Third Reich*, 784-796. New York: Praeger.

Blass, R. 2002. "Manipulation in the Speeches and Writings of Hitler and the NSDAP". Paper delivered at the International Conference on Manipulation in the Totalitarian Regimes of the XXth Century, Ascona, Switzerland, Sept. 29th – Oct. 3rd, 2002.

Bork, S. 1970. *Missbrauch der Sprache: Tendenzen nationalsozialistischer Sprachregelung*. München: Francke.

Carston, R. 1998. "Negation, 'Presupposition' and the Semantics / Pragmatics Distinction". *Journal of Linguistics* 34: 309-350.

Carston, R. 2002. *Thoughts and Utterances*. Oxford: Blackwell.

Choi, I., Nisbett, R. and Smith E. 1997. "Culture, category salience, and inductive reasoning". *Cognition* 65-1: 15-32.

Chomsky, N. 1989. *Necessary Illusions: Thought Control in Democratic Societies*. Boston MA: South End Press.

Dennett, D. 1989. *The Intentional Stance*. Cambridge, MA: MIT Press.

Eemeren, F. van and Grootendorst R. 1992. *Argumentation, Communication and Fallacies. A Pragma-Dialectical Perspective*. Hillsdale: Lawrence Hellbaum.

Glunk, R. 1966. "Erfolg und Misserfolg der Nationalsozialistischen Sprachlenkung". *Zeitschrift für Deutsche Sprache* 22: 57-73.

Grice, P. 1975. "Logic and Conversation". In P. Cole and J.-L. Morgan (eds), *Speech Acts* [Syntax and Semantics 3], 41-58. New York, Academic Press.

Herman, E. S. and Chomsky, N. 1988. *Manufacturing Consent: The Political Economy of the Mass Media*. New York: Pantheon. 1994 edition: London: Vintage.

Jaccard, J.-P. *et al*. 2003. *Un "mensonge déconcertant", la Russie au XXe siècle*. Paris: L'Harmattan.

Johnson, P., Grazioli, S., Jamal, K. and Berryman, G. 2001. "Detecting deception: adversarial problem solving in a low base-rate world". *Cognitive Science* 25 (3): 355-392.

Klemperer, V. 1946/1975. *LTI – Notizbuch Eines Philologen*. Leipzig: Reclam Verlag.

Kraus, K. 1933/1952. *Die Dritte Walpurgisnacht*. München: Kösel.

Lifton, R. J. 1961/1989. *Thought Reform and the Psychology of Totalism*. Chapel Hill: University of North Carolina Press.

Mankeltow, K. and Over, D. 1990. *Inference and Understanding: A Philosophical and Psychological Perspective*. London and New York: Routledge.

Paechter, H. 1944. *Nazi-Deutsch: A Glossary of Contemporary German*. New York, Ungar.

Rigotti, E. and A. Rocci (in press). "Congruity, Connective Predicates and Theme-Rheme Articulation". In P. Schulz (ed), *Semiotics and Communication Sciences*, University of Toronto Press.

Sauer, W. 1978. *Der Sprachgebrauch von Nationalsozialisten vor 1933*. Hamburg: Buske.

Seidel, E. and Seidel-Slotty, I. 1961. *Sprachwandel im Dritten Reich*. Halle: Verlag Sprache und Literatur.

Shmelev, A. 2002. "Manipulation in Soviet Discourse: Universal and Language-Specific". Paper delivered at the International Conference on Manipulation in the Totalitarian Regimes of the XXth Century, Ascona, Switzerland, Sept. 29th – Oct. 3rd, 2002.

Sperber, D. 1982. "Apparently irrational beliefs". In S. Lukes and M. Hollis (eds), *Rationality and Relativism*, 149-180. Oxford: Blackwell.

Sperber, D. 1985. *On Anthropological Knowledge*. Cambridge: Cambridge University Press.

Sperber, D. and Wilson, D. 1995. *Relevance. Communication and cognition* (2nd edition). Oxford: Blackwell.

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

Sperber, D. (ed) 2000. *Metarepresentations*. Oxford, Oxford University Press.

Taillard, M. 2000. "Persuasive communication: the case of marketing". *UCL Working papers in linguistics* 12: 145-174.

Wilke, J. (ed) 1998. *Propaganda in the 20th Century. Contributions to its History*. Cresskill: Hampton Press.

Zweel, H. van. 2002. "Managing consent during the Apartheid era: a linguistic perspective". Paper delivered at the International Conference on Manipulation in the Totalitarian Regimes of the XXth Century, Ascona, Switzerland, Sept. 29th – Oct. 3rd, 2002.

---

[1] First published in 1986.

[2] One could suggest that in cases of brainwashing or 're-education', sincere consent is gained through brutal force. I hold that it is far from clear that such consent is really *sincere* in the full sense of this term. But even if it were, this kind of manipulation is not the standard one, which is about someone coming to some conclusions

believing he does so freely. Manipulation seems much more efficient when the constraints are not heavy and brutal but smaller and milder, as when the Chinese gained consent of American prisoners step by step, first asking them without heavy pressure to explain about some bad things in America, then having them criticize more and more heavily their county, the process leading the prisoner to finally sincerly think that Communism is indeed better than Capitalism. They used a cognitive feature known as *cognitive dissonance* (cf. Lifton 1961/1989).

[3] Actually the *set of facts manifest to the hearer*, which presents slight – and irrelevant at this point – differences with the "set of beliefs" of the hearer. See Sperber and Wilson 1995 section 1.8 for details.

[4] This is not a definition for 'communication of truth-functionally defective propositions' but for manipulation, since communication does not imply that the communicated material becomes part of the beliefs entertained by the hearer, an implication *manipulation* has (as well as persuasion, but then the nature of the propositions need not be defective and the set of strategies would be different; I see manipulation as a sub-type of persuasion).

[5] Ideally, I should then have replaced the symbol of implication by a symbol for 'assumed by the speaker to imply...', but for the readability of this definition, I chose to simlpy add this precision in words.

[6] I use *communicative strategy* in the simplest meaning, without reference to psycho-social strategies.

[7] Explicatures are context-dependant inferences which are part of what the utterance explicitly communicates (roughly, the Gricean "what is said", although this notion is not operative, according to Relevance theory). See Sperber and Wilson 1995 section 4 and Carston (2002) for details on the explicature / implicature distinction.

[8] For obvious reasons: any set of truth-functionally interdependent assumptions is globally false if one of them is false (*P and Q* is true only when both *P* and *Q* are true).

[9] I follow Carston (1998) and our previous work on presupposition in negative utterances (Saussure 2000) in assuming that presuppositions are types of implicatures. The equivalence between presuppositions and implicatures is also assumed in other paradigms, for example in *congruity theory*, where implicatures are dealt with in terms of specific presuppositions. See Rigotti and Rocci (in press).

[10] I call *manipulator* any manipulative instance, be it a single speaker or a collective entity (the *Party* for example).

[11] A fact may be manifest to an individual although he may not be aware of it. For example, one may have shortly noticed elements of the situation of speech that are irrelevant to him, and that have not been processed as information but are stored anyway in the mind. Moreover, it is likely that not only facts can be unconsciously manifest to an individual, but meta-propositions such as *P may be false*. Being unconsciously aware of a proposition or of a fact does not imply plain insincerity: the division between *sincere – insincere* is obviously too rough for a fine-grained analysis (as the distinction *lie – not-lie* may also be). In a relevance-theoretic terminology, we would say in this very case that *the fact that the speaker's belief is false or doubtful is manifest to himself*: a fact is manifest to an individual if this fact is known or *could be known without any further information* (a fact is manifest if known or inferable).

[12] The bandwagon fallacy is a non-formal fallacy that can be expressed as follows:

P: Everybody / most people / a great deal of people believe Q

        Therefore, Q is true.

     Arguments of authority are of a similar form (and is the same in essence):

        P: This great man believes Q

        Therefore, Q is true.

[13] Utterances and propositions are distinct things: we reason with many un-uttered propositions, and the interpretation of a single utterance can lead to several propositions as meaning.

[14] Since the emergence of the National-Socialist party, intellectuals pointed out a few commonsensical characteristics of manipulative discourse in the public speeches of the party, for instance the "political smoke-screen of phrases" pointing out fuzziness (see in particular Kraus 1933 / 1952) and many recurrent elements such as slogans and proverbial manipulation. The major work, although mostly descriptive and not theoretically detailed, is Klemperer's (1946 / 1975 / 1998). He addresses a wide range of issues and labelled the way the party was using language *Third Reich language* recalling Orwell's Newspeak. A heavy bibliography on philological issues is available, from which some must be quoted: Berning (1964), Betz (1955), Bork (1970), Glunk (1966), Paechter (1944), Sauer (1978), Seidel and Seidel-Slotty (1961).

[15] I assume a distinction between 'pseudo-mystical' discourse and religious discourse in general. Religious discourse is not manipulative in essence, since it may well be provided in good faith (it can be persuasive, which is another matter); it may become manipulative when produced by religious extremists in order to promote some political ideology and some individuals' power. In this case, we suggest an internal contradiction between the belief in a 'true' God, symbolising the ultimate 'absolute', and the authority of a human leader claiming to hold himself the absolute truth about God. Since God is normally *out of the natural ontology*, he always remains "unknowable" in some way. Claiming to hold the full and only truth about what God thinks and wants is therefore manipulative because it is contradictory to the notion of God itself, at least within monotheist traditions. We notice that the preachers presenting themselves as holding such a full and only truth make an extensive use of manipulative strategies like appeal to emotion and arguments of authority.

[16] I do not put prosodic features and intonations into linguistic aspects of communication but into pragmatic and contextual ones; I am aware that this is a controversial issue.

[17] The German *ent-* corresponds, roughly, to the English *un-*, *de-* or *dis-* in newly created words like *Entdunkeln* ('undarking', the operation, taking place in the morning, consisting in taking off the screens put on windows to darken the city when an allied bombing was likely to happen), *Entrümpeln* ('disobstructing' the attics for easy access in case of fire) etc. Klemperer (1946/1975), who reports these expressions, notes and regrets a similar word-formation in *Entnazifizierung* (denazification) after the fall of the Third Reich. One notices also the word *destalinizatsja* in Russian, built up under the same scheme ('destalinisation').

[18] The assimilation of sport and military activity probably had other motivations in the antique world, when there were no guns, bombs and planes, and where the physical ability of the soldier was the key to victory. The remotivation of this parallel between health and war had a great importance in the Third Reich, probably because it was intended to reinforce the idea of the superiority of the pure Arian race over inferior (and therefore weaker) peoples. But it also links to the Nazi mythology of antique, legendary and heroic germanity, where nature (and therefore the body) plays a particular, quasi-mystical role. In some democratic countries, there are still remains

of these analogies – see for instance the strange Swiss Federal Department (ministry) of *Defence, Protection of the Population and Sports*.

[19] This is not a law. To take an example, if a doctor starts saying weird things, a patient may suspect incompetence, and, as a result, confidence is lost. However, *a priori*, it is legitimate and rational to suppose that doctors deserves confidence – otherwise, it would be of no use consulting them. Another comment on this further down.

[20] Speech in Buenos Aires on October 17th, 1943.

[21] One might talk about the lack of global *coherence* of the manipulative discourse, but this notion is too intuitive and problematic to be used within our framework; see Rocci (in this volume) for developments on manipulation and coherence.

[22] Simplification has other interesting consequences. In general, a public discourse relying on simplification has more power of convincing than nuanced and more elaborated discourses. This is a side-effect of the global economy of human cognition: a high level of political expectation raised by simplistic discourse for a minimal reasoning, a minimal cost. In order to adhere to a more complex discourse (which represents a more complex and interrelated set of propositions), one has first to foresee a *more adequate* political result (in terms of justice, efficiency, etc.) that justifies a wider set of elements to deal with – and therefore to accept a longer and more complex information processing and reasoning.

[23] I am aware of the complex relations of religious faith with non-religious, ideological, dogmatic beliefs. I hold that there is a kind of religious discourse that is indeed totalitarian, ideological and manipulative, and another kind which is not, and that these two kinds of religious discourse have occurred in most religions (see note 15).