

UCLA

Other Recent Work

Title

The Chinese Room's Secret

Permalink

<https://escholarship.org/uc/item/1mw1b584>

Journal

L'Homme, 150

Author

Jorion, Paul

Publication Date

1999

Peer reviewed

The Chinese room's secret

Paul Jorion

Abstract

Starting from a reflection by Jean Pouillon, it is shown - both deductively and on the basis of experimental data - that consciousness is deprived of any decisional power. Consciousness' role is reduced to transmitting to the body instructions based on the emotional response to percepts. Language allows human beings to generate a self-justifying narrative of their deeds. Such an account does not reflect, however, the actual psychological mechanism at work. Consciousness' actual effectiveness resides in influencing on the one hand the affect of the speaker (as speech or as "inner speech"), and on the other hand the affect of any listener. The pair "body" and "soul" gets thus validated, but their traditionally assigned responsibilities need reassigning between one body that decides and acts and one soul whose feedback operates at the affective level only.

Keywords

consciousness, Benjamin Libet, willpower

The paper is the English translation of an article initially published in the French anthropological journal *L'Homme*, 150, 1999: 177-202. At the time of writing the author was Regents' Lecturer at UC Irvine.

The 1999 French text can be found online as part of the Persée archive: [Le secret de la chambre chinoise](http://www.persee.fr/doc/hom_0439-4216_1999_num_39_150_453573) : http://www.persee.fr/doc/hom_0439-4216_1999_num_39_150_453573

The first draft of "Jean Pouillon and the mystery of the Chinese room" was very different from the article published in *L'Homme* in July 1997 (Jorion 1997b). In the original text, I carried out much more faithfully the task I had accepted: to account for Pouillon's contribution to anthropological theory.

Developing an argument presented for the first time in my account of *Le cru et le su* (Jorion 1994b), I summarised Pouillon's point of view as "the sociology of *modus vivendi*," based on the idea that men's life is arduous and leads them to be content with the tolerable minimum: In the best of cases, politics thus gradually comes to acquire the transparency of the crystal desired by

Rousseau; more often than not, unfortunately, the effort of rationalization is interrupted very early on and the abundance of individual strategies generates an opacity of the polity, similar to the sombre view established by Hobbes.

The only link between this initial version of my text and what later became “Jean Pouillon and the mystery of the Chinese chamber” lies in this sentence: “Awareness only occasionally comes into play in the lives of individuals, such as the need for a cultural invention or progress in ways of explanation, in that of societies.”

From rewriting to rewriting, the text metamorphosed into a discussion of a Pouillon article originally published in 1987, “Le plaisir de ne pas comprendre.” As a reminder, here is the passage to which I devoted my comment:

For more than thirty years, I have been making a summary report of the meetings of the ECSC Consultative Committee, in which producers, users and workers of coal and steel have been engaged in highly technical discussions wherein I have not understood anything and have never sought to understand anything. (...) I don't know the meaning of the words they use - what are, for example, coils, long profiles, quarto plates? Their allusions to manufacturing processes or economic problems remain as mysterious, but less exciting, than the stories my grandmother once told me. This ignorance is by no means a handicap: it is exceptional that the record of a debate was subsequently contested by those who participated in it. They understand what I wrote without understanding. Of course, I could learn the meaning of the terms: the (relative) pleasure I take on this task would then vanish. It consists in forging sentences, often more “correct” than those used by the speakers and whose grammatical structure guarantees me that they can have a meaning. Which one? I don't know, but it must be there since others than me can find it. (...) Thus form carries away the substance, and one can devote oneself to it in a manic manner, without harm to it; if language is a means of communication, one can hear it function by remaining outside communication, within that margin where the hearer-viewer stands, on the threshold which he does not wish to cross (Pouillon 1993:155-157, Jorion 1997b: 92).

The question was thus admirably laid out by Jean Pouillon of the relationship between the individual words' significance and the global meaning of the sentence, an issue that scholastics had debated – but failed to solve – under the name of *complexe significabile*.

Considerations made by Pouillon earlier in *Temps et roman* (1946) allowed me to complete the examination of the question, and progressing in a deductive way, I ended up with the statement of three theses, logical consequences of the quoted text, but of which I saw myself obliged to immediately take up the flagrant absurdity of the last two:

1° knowledge of syntax rules is essential for the correct use of a language,
2° knowledge of semantic rules is indifferent to the correct use of this language,

3° the intuitive sense of comprehension is also indifferent to the correct use of a language (Jorion 1997b: 96).

The article ended with an admission of failure. Locating the heart of the debate in Jean Buridan's question about the meaning of *syncategorems* - the words providing the syntactic framework of the sentence - I was writing: I have not solved here the question of the meaning of the syncategorems that Jean Buridan called "the scourge of logic," not "for lack of space" according to the admitted phrase, but more simply because I do not know the answer "(ibid. 97).

* * *

The parable of the Chinese room was originally conceived by the philosopher John Searle: "Imagine that you are locked in a room, and that in this room are various baskets filled with Chinese symbols..." (Searle 1984:32). The mystery of the Chinese room is this: having spent a certain number of days or years in the room, you end up behaving "exactly as if you understood Chinese," the fact remains however that you do not understand a single word of Chinese" (ibid. 33).

A double recent experience has taken me back to the troubling precincts of the chamber, and the re-examination of its mystery has allowed me - or so it seems - to clarify it. First of all, I read once again my own text a year after it had been written, this time in the tribute volume of *L'Homme* for Pouillon. On the other hand, and at the same time, I wrote - again at the invitation of *L'Homme* - an account of the *Dictionnaire de la psychanalyse* by Elisabeth Roudinesco and Michel Plon (Jorion 1998). It is this conjunction of reflections relating, on the one hand, to the meaning of the sentence and, on the other hand, to *anamnesis*, that is, the process of recollection in the psychoanalytic cure, that has allowed me - I believe - to uncover the secret of the Chinese room.

In Poe's *The Purloined Letter*, the Minister's stratagem consists of not hiding the stolen missive, but on the contrary, putting it prominently on the mantle of his chimney having undergone but a minimal transformation: reversing its folds. The Minister counts, of course, without Dupin's insight. In the case of the Chinese room, the solution to the mystery is also at hand, the trick - as we will see - consisted here for the master of the game to place the most insurmountable obstacle between us and the answer: man's vanity.

* * *

From the 1960s onwards, Benjamin Libet, a neurophysiologist at the University of California at San Francisco, undertook a series of experiments on consciousness, one of the possible - and certainly most plausible - conclusions of which is that it has no causal power. Consciousness is a window open to the

world, a “glance,” but any power that one might try to assign it, as the place wherefrom decisions are actually taken, seems illusory. In other words, on what would be called the “rational” level, where a faculty such as “will” could be exercised, consciousness is but a dead end whereto information can no doubt reach, but without there being any back process of a decision-making type. It is at the level of the *affect* (emotion), and of it alone, that the information displayed in the aperture of consciousness produces feedback but of an “involuntary,” automatic nature.

Other interpretations of the results of Libet's experiments are undoubtedly possible, some of which he himself has stated over the years, essentially under pressure from the criticism he has been subjected to, all aimed at the same goal: to offer an alternative to the conclusion I mentioned, which seemed universally inadmissible in the world of what is nowadays labelled “cognitive sciences” and used to be called psychology. Thus, in his book *Consciousness Explained*, Dennett describes the thesis of the absence of causal power of consciousness as “incoherent” (Dennett 1991:164), without however providing any justification for the fact that he characterises it in such terms.

Libet proposed, as other hypotheses compatible with the results it leads to, that consciousness exerts its influence through time travel, or that it is of a field effect nature, but without this field being detectable “by any objective, physical measuring instrument” (1997:137), otherwise it would have already been recorded by one of the very many devices invented for such purposes.

Other researchers, such as recently the philosopher David Chalmers (1996), have suggested that consciousness is a phenomenon whose nature is perhaps radically different from that of other manifestations of the physical world: namely, immaterial, and metaphysical.

It is noteworthy, however, that none of these three alternatives to the hypothesis of the causal impotence of consciousness is a classic explanation in science: in the realm of science, as we know, physical effects are always detectable. As for metaphysical forces, they exclude themselves by definition from the field of physics and, being unable to belong to any other field of science than this one, they are thereby excluded from its field of application. Only one of the three, the hypothesis of time travelled backwards, has a semblance of legitimacy, having been mentioned in nuclear physics - in particular by Richard Feynman - about the behaviour of certain anti-particles.

Of the four competing hypotheses, only that proposed by Libet of a causal impotence of consciousness, thus presents the usual features of scientific explanation. Its unanimous rejection by psychologists in the name of its obvious “absurdity” is actually a simple consequence of the fact that this thesis is in fact at odds with the epistemological credo of “cognitive sciences,” namely the “functionalist” scheme borrowed from Brentano. This scheme can be briefly characterised as follows: Consciousness is the ability to concretise an intention

into the intended act; an intention is determined by a desire based itself on a justified belief, desires and beliefs are mental states corresponding to specific material configurations of brain cells (Searle 1997:44).

It should be noted in passing that, with the exception of the mention of neurons, of relatively recent discovery, this functionalist scheme is identical to “popular psychology” as it is inscribed in current language since ancient Greece. By contrast, the theories relating to the functioning of the human mind - starting with those of Plato and Aristotle - have always challenged the *obvious* embedded in the way language predisposes us to talk about things. Brentano's system thus constitutes a return to the spontaneous apprehension of mental functions as suggested by language and thus deliberately navigates in the rough waters of “epistemological naivety.”

The causal role ascribed by functionalism to consciousness is crucial, since consciousness and consciousness alone has the power to transform an intention into an act: denying it this function amounts to depriving an origin for “voluntary” acts altogether, and thus amounts to destroying the entire conceptual edifice. Hence the so-called “obvious absurdity” of the hypothesis of impotent consciousness, at least within the dominant paradigm of the so-called “cognitive sciences”: if consciousness is impotent, the functionalist scheme evaporates.

There is another discipline, however, for which, without belonging to its classical or orthodox form, the thesis of the causal impotence of consciousness appears a priori acceptable. This other discipline is psychoanalysis.

It is a fundamental theme of Freudian metapsychology that consciousness deceives itself when it assumes that it is fully and entirely determined by itself: acts of consciousness are at least partially defined by an external influence, this one of unconscious origin. The instance of the Ego, initially posed by Freud as almost identical to consciousness, sees itself sliding over the years on the side of the unconscious, until it finds itself straddling on either side of the conscious/unconscious separation, because of what appears then as a “splitting.”

Denying any causal power to consciousness automatically means assigning the entire determination of human behaviour to the unconscious. Or, reformulating the same proposal in such a way as to avoid *substantiating* any instances such as the unconscious: any decision-making process is of unconscious origin. Although this new thesis is more easily assimilated to the theoretical corpus of psychoanalysis than to that of psychology, it is at the moment foreign to it. For Freud, consciousness has an obvious decision-making role: to be the place of the first observation of perceptive data, which are then transmitted to purely cognitive devices that will ensure their processing.

However, Libet's observations are formal in this respect: perceptive data only reach consciousness about half a second after the event has been perceived, a

period of time during which so-called “voluntary” motor acts in response to these percepts may have taken place with much shorter response times of the order of one-tenth of a second. This means that the delay with which consciousness is informed of a possible reaction of the body to percepts - in case of immediate danger, for example - is about four tenths of a second. Freud's supposed direction in transmitting information provided by perception, i.e. from the conscious - in direct contact with perception - to the unconscious, is thus invalidated by experience, and the results obtained by Libet force a revision of a central scheme to Freudian metapsychology.

Consequently, the most plausible conclusion that can be drawn from the work of the American neurophysiologist, if it does not appear a priori as heretical in relation to psychoanalysis, nevertheless obliges a reworking of the theory such as formulated by Sigmund Freud. Lacanian metapsychology, built as claimed by its author on the foundations of the Freudian edifice, is better able to assimilate without major modification the thesis of consciousness as a decisional cul-de-sac; insofar as in Lacan, the *subject* of the Unconscious is an effect of the signifying chain, any decision-making entity in individual *subjects* automatically finds itself at the end of the command line.

* * *

Regardless of how much credit can be given to Libet's work, we must find answers to the questions of common sense that his discoveries necessarily raise, and in particular, if it is true that consciousness is deprived of causal power, why does it seem to have such power? if its function is not of a decision-making nature, what was its aetiology? Finally, if its apparent function is illusory, what is the actual function of consciousness?

Why does consciousness seem to have causal power?

Consciousness seems to have a causal or decision-making power because of an effect that Libet calls “backward referral,” a term that I will translate, respecting the spirit more than the letter, by retrospective assignment. This effect, verified experimentally by him, ensures that any percept - no matter how long it takes to be interpreted by the nervous system - is resituated in the consciousness rebuilt chronology at the exact moment when the perceived phenomenon was actually recorded: “Such subjective referral serves to “correct” the temporal distortion of the real sensory event, a distortion imposed by the cerebral requirements of a neural delay for the experience.”(Libet 1992:267).

If it is indeed proven that 500 milliseconds pass before the information obtained by the senses is displayed in consciousness, it follows that this faculty is powerless to exercise the decision-making function implied in the “voluntary” acts that constitute an immediate response to this information. Indeed, triggering so-called “voluntary” acts requires only a much shorter response time compared to percepts: Libet observes, “behavioural responses to sensory signals or images can be made within as little as 100 milliseconds” (Libet

1992:267).

The neurophysiologist's first works to ensure his renown were those wherein he highlighted - in my opinion, incontrovertibly - a delay in the subjective experience of formulating an intention in relation to the actual performance of the act which that intention is intended to achieve, in other words and in order to clear any ambiguity, Libet highlighted that the feeling of initiating an act actually comes after the actual act itself, and this with a delay of the order of half a second. In the author's words: "the brain has begun the specific preparatory processes for the voluntary act well before the subject is even aware of any wish or intention to act" (ibid. 263).

What then leads a subject to assume - wrongly - that the intention precedes or is at least simultaneous with the "voluntary" act? The very same mechanism of retrospective assignment. It ensures that the awareness of an act taken by a subject seems to coincide with its realization, and can - as a result - be conceptualised or "rationalised" as having been, not a passive awareness, but in fact the active intention that was realised in this act.

The term "voluntary" of course reflects the fact that these acts are traditionally attributed to the exercise of willpower, a classic denomination of the decision-making power of consciousness. The underlining of the retrospective assignation and the delay of the formulation of the intention in relation to the achieved action prohibits postulating the existence of a function such as the will, which would prove having been historically yet another "phlogiston." As a matter of methodological principle I exclude the suggestion that a decision-making function can be exercised back in time.

The logical consequence of the above is that any decision of positing an act - whether it is traditionally described as voluntary or involuntary - is in fact unconscious, and that any representation of the completed act in terms of an intention which was carried out is necessarily an ex-post interpretation and of the nature of a rationalisation.

If the function of consciousness is not a decision-making one, what was its aetiology?

The retrospective assignation effect was only experimentally corroborated by Libet for the sense of touch. However, he considered it highly likely that a similar phenomenon would occur for each of the senses and therefore for each type of perception. Such a mechanism, which would take into account the different processing times of the data captured by the five senses, would allow a subject to see on display simultaneously within consciousness the percepts that were actually recorded simultaneously by each of the senses, about half a second beforehand.

The five specialised senses of mammals allow them to simultaneously experience sensations of distinct orders: luminous (electromagnetic), chemical,

mechanical and acoustic. Associated with a data of internal origin: the *affect*, these configurations of sensations perceived simultaneously constitute, according to the expression introduced by Jung and Riklin (Jung [1906] 1973), “complexes.” The role, which then falls to retrospective assignment, is to ensure that these sensations are authentically present simultaneously in active memory. These complexes are part of long-term memory as « remembrances, and together form what is called the “memory” of a subject. Hence the critical importance of Libet's discovery: retrospective assignment allows the recording under the shape of a single memory trace of stimuli of various perceptual origins which were truly synchronous from the subject's standpoint, as well as an *affect value* [emotional weight] constituting per se a signal directed to the body.

The *affect* is the emotional response that corresponds to each of these complexes. This is what consciousness perceives of hormonal instructions at the origin of motor acts, whether it is their preparatory stage, their work in progress or the residues of such signals. Gazzaniga and LeDoux have recently resurrected a very similar concept, a legacy of William James who considered emotions as interpretations by consciousness of hormonal manifestations of the body (LeDoux 1996: 43-45).

The meaning of the “present moment” would correspond in this view to the chronological thickness of the “aperture” to which consciousness is identified, i. e. the period of time within which percepts will be considered - for the purpose of recording in long-term memory - to have been synchronous; that is, the “instantaneous” lifespan of the active memory.

If its apparent function is illusory, what is the true function of consciousness?

Within our memory, remembrances are linked together through those feelings they share. It is the existence of this network of connections that makes it possible to recall: the evocation of a memory from one of the emotions that compose it. A memory can thus be represented (in the double sense of “reappearing” and “being the object of a representation”) in its entirety, i.e. as the configuration of feelings that have been perceived simultaneously.

Thus, every feeling has a double capacity: that of inscribing itself within memory as a component of a remembrance - which is none other than a “complex” of sensations which were simultaneous - and that of evoking on a virtual stage which I will call according to the use, *imagination*, ancient memories whereof it is one of the components. The storage of remembrances in memory thus allows a feeling experienced at a given moment to generate later in the imagination of a subject a delayed representation of all its previous avatars.

The real function of consciousness then lies at this level alone, in the recording of simultaneous percepts in the form of memories and their subsequent evocation - associated with an emotional weight - when similar sensations

present themselves once again as percepts. Paradoxically, processes that find their point of origin at the level of consciousness are therefore emotional and not rational. We will see that the half-second lag between perception and awareness plays a crucial role in the efficiency of the mechanism: it is this very lag that allows a fluid self-correction of behaviour, and is the source of the illusory representation of a decision-making function entrusted to consciousness.

The ability - within the *aperture* of consciousness - to record memories made up of synchronous percepts, has a clearly adaptive function: such reconstitution of simultaneity within memory is the crucial condition for true learning, that is to say the assurance that in the future signals of the same type coming from the outside world will trigger a reaction integrating all the information accumulated by a subject during his history. When recalling a sensation in reminiscence, the remembrances whose one of the components belongs to this sensation will be represented in their unique combination of feelings and emotional weights, constituting through their combination what could be called an affective mood. The eminently Pavlovian character of this mechanism It will not have escaped the reader.

Libet, on the other hand, emphasizes the adaptive role of the subject's decision-making power being unconscious rather than conscious in its entirety: "Given a neural delay of several hundred of milliseconds for developing conscious awareness of a mental event," he writes, "it is clearly of great adaptive value that much or most of our mental processes proceed unconsciously, without the lengthier neural requirements for awareness having to take place" (Libet 1992: 268). It will have been noted that the neurophysiologist writes, "much or most of our mental processes proceed unconsciously," whereas the conclusions of his experimentation cover at first glance all possible cases. We will see, later on that the problem arises in the exact same terms when it comes to internal stimuli, as it does for speech and thought envisaged as "inner speech."

* * *

This reversal of perspective on the role played by consciousness is of course radical. The necessary reworking of the representations is similar to that which captures the last survivor of the crew in *2001: A Space Odyssey*, when he suddenly realizes that the mission was actually entrusted to the spacecraft and its obsessive computer, not to the human crew - probably judged by the mission designers to be less reliable. As a result, many common phenomena, about which words of the common language offer us a spontaneous psychological interpretation, must be reinterpreted with a view to their actual operation: we need to reallocate their actual responsibilities to mental functions, and determine in particular what are the causes and effects. Not all is illusion, though: consciousness as such is not an illusion, but its decision-making power is one; my intention to pronounce a sentence at the moment I utter it is illusory, but my intention to go to Santa Monica, on Tuesday 4th, is

very real. And so on.

To facilitate such a transpose, I will - like Confucius - suggest a “rectification of the names.” I used a few paragraphs ago, the word imagination to designate the “virtual scene” whereon reminiscence takes place. Since it is on this same scene that percepts are displayed - with the delay mentioned above of half a second - there would be no great shame in writing “imagination” where I used to write “consciousness,” and that is what I intend to do most of the time. Likewise, I will write - without further justification - “body,” where I used to write until now “unconscious.” We will now say, names having been corrected: “In truth, decision-making, the will, has been entrusted to the body and not to the imagination.”

The major advantage I see in this change of vocabulary is that it makes it easier to escape from certain habits acquired with psychoanalysis: those, for example, of an unconscious whose role is limited to introducing distortions in the field of consciousness, similar to the effect that the magnet moved under the sheet of paper at school had on iron filings. Consciousness must henceforth be conceived as a dead end, the word *imagination*, with its own connotation of evoking a world of fiction, makes things easier.

* * *

We must learn to rewrite in a different way all scenarios where consciousness appeared in its traditional role of conductor of events. A few years ago, in *Principles of Intelligent Systems* (1990), I had already introduced a model for thinking avoiding the pitfalls of explanations invoking *final causes* and finalities.

A drop of rain falls at the top of a mountain. Little by little, day-by-day, it makes its way to the ocean. At first it travels in one hour several kilometres - playing with every obstacle - then its speed slows down at times approaching immobility. It will sometimes stay for many months in the same water body, a lake for example. Yet - unless it evaporates on its way - it will inevitably one day reach the ocean. All raindrops over the ages have been moved by the same infallible instinct allowing them in spite of every hardship to reach the sea.

Such a way of presenting things is, of course, ludicrous: raindrops are by no means possessed by an instinct; nor is there any deliberation on their itinerary: whatever the apparent complexity of their route, it is determined by a unique physical principle, that of gravity, which means that at every point and at every moment, the drop adopts the direction indicated by the steepest slope. At a bird's view it may seem to skilfully jump over obstacles, hesitate about the direction to take, seem to rest, and so on, but these appearances, we know, are deceptive: the drop follows the gradient of the steepest slope, full stop. The tiny runoff will make the broad river and at the instant when the small drop of rain that first trickled on the ground meets the ocean, all will have happened as if the sea had called it, from that moment in the future when the junction is

finally taking place. However, the ocean just needed to be there: the ocean is simply what we find at the end of a gradient, what we call in physics, a potential well.

Does this gradient model make it possible to replace the classical model of human behaviour in terms of intentionality? For sure: it allows us to dispense with any notion of a forward projection, any tension which would guide us towards a goal (Jorion 1990: 91-96; 1994a: 94-98; 1997a: 3-4). If it is a simple movement along the steepest slope leading to relaxation when a potential well is reached, this suffices to deprive consciousness of the most ambitious function that we usually assign to it: to deliberately move towards the goals that it has previously set itself, and this - let us note - quite independently of any experimental evidence such as that provided by Libet.

When the shopper says that he or she has “my glance » drawn to a particular store’s window on the shopping street, his or her way of expressing himself or herself very accurately reflects the mechanism at work. The subject, his memory, as well as his physical environment, together constitute a universe of possibilities, within which his behaviour draws a path towards the lowest point, the one corresponding to his satisfaction, to the temporary cancellation of a felt lack. The expression « my glance being drawn “ establishes the observation of the existence of such gradient’s logic where we, human subjects, allow ourselves to get captured - as the onlooker is captured mechanically by the windows on the street where he wanders - by the steepest slope of the configuration space that constitutes the double complementary system of ourselves and the world surrounding us.

* * *

Man's originality compared to other animals lies in his use of words in speech and thought. The word spoken in speech is a sensation perceived by the ear, the written word is a sensation perceived by the eye. Like any sensation, the word is likely to be experienced simultaneously with others and therefore stored in memory as a remembrance. Similarly, like any sensation, a word can evoke in a different way in the imagination other sensations whereto it is associated in particular remembrances. The sequencing of words and sensations of perceptual origin associated with them through recalling them in the imagination is one of the constituent elements of what we call thought.

Words are exchanged by humans among themselves. A word uttered by one elicits reminiscence within another. In pronouncing a word, or in writing it, a subject automatically imposes on another subject to whom he speaks, or to whom he writes, to link in his own imagination the memories with which these words are associated for him (Ockham and Guillaume of Saxony hold that “a sign is all that can remind a person of something he already knows” - Moody 1953:18). Thus by dialogue, by exchanging words, men are able to evoke each other's memories. Better still, since their conversations are in themselves

matters of remembrance, new ones that will be common to them constitute themselves through their dialogue: in other words, through dialogue they build a common part of each other's memory.

Dialogue is therefore an automatic process: the evocation of memory by one of the interlocutors generates the reminiscence of the other. The conversation is exhausted and then extinguished when the memory of one person no longer calls for reminiscence in the other. If they have the freedom to do so, then they leave. But if the time of one belongs to the other because of a particular balance of power between them, then the former is obliged to remain and his prolonged presence condemns him to store in his own memory what was until then only the memories of the latter. His memory was annexed by the memory of another and is now an extension of it: according to the expression enshrined in philosophy, he is "alienated."

The Greeks were the first to realise that one could impose one's memory on others and thus alienate them without resorting to force, without even taking advantage of any pre-existing balance of power. The way memories are connected allows remembrance to travel through them in multiple paths, but if the speaker evokes memories in a special order, then those things he knows will automatically become the knowledge of the listener. Such magic, which allows someone to impose his memory on another person without making the latter feel violated, is conviction; the Greeks compared this power to that of a drug which alters the state of consciousness, and called it *pharmakon*. Aristotle was the first to codify the order in which his or her memories should be evoked to impose them on the memory of others, to "win his heart," to carry his conviction away. The way of proceeding if we start from certain principles, he called it *analytical*; if we start from simply plausible principles, he called it *dialectics*.

The method according to which words are uttered in a way that is not a pure and simple "bulk" remembrance - what I have just called "evoking one's memories in a special order" - is what is usually called "reason," and it is with simple reminiscence, the second mode of thinking. To evoke one's thought in a precise order is to reason, and the memory trace linking two others, is the *reason* linking the final one to the first one (classically, in Aristotle's system, *reason* is the *middle term* of the syllogism, the one vanishing in the conclusion to leave solely present the two *extremes*, between which there is now a relationship that did not pre-exist to the syllogism; in ancient mathematics, the *reason*, of course, is the arithmetic or geometric *mean* - between two terms; see Jorion 1996: 281).

In the new perspective opened up here, this evocation in a precise order is however not deliberate, it follows a gradient which is in this case constrained by the ordering that Aristotle codified under the name of *sylogism*. It is depending on whether thought leads to one location or another in memory traces that following the slope will be generating syllogisms or « simple connections » - akin

then to free association (Jorion 1990a: 52-54). It is the culture within which anyone of us is embedded that proposes some fields where notions are nested by a relation of inclusion of one another in each other and some other fields where they are connected only through identity of nature, similarity or proximity in time and space (Jorion 1989).

The absolute core of each human subject is what Lacanian metapsychology defines as his or her “fantasme”: the “filter” constituted by the whole of what I have called in *Principles of Intelligent Systems*, the “nuclei of belief” (Jorion 1990a: 64, 66,138), the primal inscriptions associated with the primal emotional weights, those that were contemporary of this inscription. When the metabolism degrades as in dying, the last accessible signifier is logically that of the first inscription: “Mother.” Nor is it surprising that the image of a tunnel at the end of which there is a light appears visually at the moment of death, an image that is no other than the reminiscence of the “primal” visual memory: a light - the first light experienced - at the end of a tunnel.

The truth of a conversation develops as the complex (fractal) boundary of the two attractive basins that the *fantasmes* of the interlocutors constitute. In a forthcoming book entitled *Le prix* [since then published in 2010] I explain how price gets formed in the exact same way as truth arises: the two phenomena can be explained in the same way because their structure is isomorphic; the only difference between them is that truth expresses itself in a *words* mode while price in a *numbers* mode. If we talk about the *truth*, we are talking about something that works as the *price* assigned to words’ exchanges between two interlocutors and if we talk about *price*, we are talking about something that works like the *truth* of human relations.

One could claim in a lapidary way that *price is the truth of human things expressed in numbers and truth, the price of human things expressed in words.*

* * *

It was Freud who first observed that without an emotional dynamics to provoke remembrance, our stored memory would only be a static network remaining without any external expression. Words have an emotional value attached to them radiating on those directly connected to them. These values are in perpetual variation: let’s imagine the fabric of memory associated with emotional values as akin to the ocean’s surface by strong wind.

When we are awake, thought travels through the network of memory traces according to a dynamics determined by the emotional values: it is emotion which channels, which decides at certain bifurcations that thought will follow this branch rather than another. In fact, the emotional values determine the *gradient*: it is them who define what is at every single passage point “the steepest slope.”

In *Principles of Intelligent Systems* I gave examples of worries that encourage,

that guide the free association of words towards them in an insistent, constantly renewed way, at any time of the day: “remember to buy butter,” “the tax reminder that was oddly immediately cancelled,” etc. (1990a: 75-76). Concern is a *potential well* about which there is nothing we can do about, it draws free association towards its bottom.

Intent can be seen as a “concern” of a particular type. As soon as the intention is present, as soon as the project exists, for example the creation of an object - and whichever the way this concern has arisen - the goal to be achieved acts as a *potential well*, even though its reality, yet virtual, is that of an outcome: the path to it is still to be built, and on a “piece by piece” basis, with the sequence of choices determined by the relief of the mnemonic landscape, and not in a single block.

Wittgenstein often questioned the nature of intention. He wonders, for example, “I have the intention of leaving tomorrow” - When do you have that intention? The whole time, or intermittently?” (Wittgenstein 1967: 10). The answer to his question is in truth “all the time in the body and intermittently in the imagination.” \

When my Santa Monica's appointment for the 4th is scheduled, some anxiety is attached to the intention of being there and being there punctually. All the words and expressions related to this event are now affected by that anxiety: the name of the place, “Santa Monica,” the name of the person, “Debbie” or “Deborah,” the date, “November 4,” the day of the week, “Tuesday,” etc. Every time that anything recalls on the stage of imagination, through association, one of these labels, the “intention” to go to this meeting will get reactivated.

Unless there are intermediate goals, intermediate tasks to be accomplished, until the said day, the anxiety felt with each evocation of the goal to be achieved will be minimal: it will be nothing more than mere *reminders* of the coming appointment.” Oh yes, I have an appointment in Santa Monica with Debbie on Tuesday, November 4th .” On the day, there will be a new association “4 November” / “today,” the level of anxiety will then be much higher and - like the little drop of rain which in each point chooses the steepest slope -, I too at each point of time and space will choose the direction most likely to reduce my worry and in the highest possible quantity.

* * *

But all of this is just common sense. Where then does the mechanism of “backward referral,” retrospective assignment, highlighted by Benjamin Libet, fit into speech? It must be assumed that in this case, too, the feeling of an intention that comes true stems from this reassignment. Measured in real time, my awareness of what I have just said must follow with the half-second delay observed in all other cases. Therefore, one cannot have in any practical way time to “intend to say” everything that is actually said in the heat of the action of a conversation or speech, i.e. once it is launched on its slope, speech

continues until extinction, *relaxation*, i.e. until the gradient of affect dies in a well of potential. In the conversation, it is the other person's discourse that, stirring my emotion, relaunches the process, i.e. creates a new gradient.

Now, and that goes without saying, one does not stop thinking when one stops talking.

Socrates: Do you call thinking what I call thinking?

Theaetetus: Tell me what it is.

Socrates: The discourse (*logos*) which the soul holds with herself about what she sees. It seems to me that when she thinks, she does nothing but talking, asking questions and giving answers, affirming and denying. When she is fixed and no longer in doubt, we call this "opinion" (*doxa*). So what I call forming an opinion is talking (*legein*) and opinion is speech that one utters not for someone else or out loud (*phonè*) but in silence for oneself.

(Plato, *Theaetetus*, 189E - 190A, based on the translation by J. Burnet, in Kneale & Kneale [1962] 1986:17-18).

You can hear yourself talking when you talk, but you can also hear yourself talking when you just think. Let's go further: if what we say we never had "the intention of saying it," then what we say, we only learn it - like anyone else - at the moment we hear ourselves say it. Merleau-Ponty once wrote,

"even I am surprised by my own words and they teach me my thought To express, for the speaking subject, is to become aware; he does not only express for others, he expresses in order that he himself learns what he is aiming at We ourselves who speak do not necessarily know what we are expressing any better than those who are listening to us (Merleau-Ponty [1951] 1960: 111,113,114).

And if we hear ourselves say what we are saying, then what we hear is just as moving to us as what we hear from others.

What does this involve? This implies the following: our discourse (both inner and outer) at the moment it is heard, modifies our affect, i.e. changes the profile of the emotional gradient underlying our speech even as it is unfolding. There is feedback, loop effect, and as with any loop effect - *cybernetic* effect - the dynamics feeds off the slight *delay* between the "hearing myself say it" and the "moving me."

"Backward referral," the *retrospective assignment*, intervenes there, in the same way as in any other type of perception, in speech and in the inner dialogue inherent to certain thought processes, generating a feedback loop allowing the emotional dynamics underlying a universe of *signifiers* stored as memory traces to self-sustain until a relaxation intervenes because of the very fact that sentences are produced, that a discourse gets generated, until it stops as

“nothing more needs to be said.”

* * *

Because the mechanism of speech is actually much closer to our conceptualizations than that of motor acts, it is the analysis of the way in which feedback for speech takes place that will enable us to understand in a new step the way wherein emotion controls behaviour.

The emotional weight stored in memory, linked to a particular memory, is a bundle of hormonal signals, i.e., a set of instructions addressed to the body in view of a readiness to act. Current percepts trigger the recollection of remembrances to which these percepts are already associated. All relevant information available to the subject regarding the existing situation is thus reactivated and, in particular, related emotional weights. The first recollection evoked may summon others to follow, according to the line of decreasing relevance that is similarly observed with *free association*.

The mood experienced is the result of the display in consciousness of hormonal residues and signals generated in response, either to current percepts or to recollections induced by them. The emotional poles of fear and anger correspond respectively to the submissive and aggressive dispositions to act. Such emotions generate by feedback and with the half-second delay observed empirically **in respect** to the recording of the percepts, a hormonal reaction which constitutes the context wherein the balance between submission and aggression is gauged, and gets materialised in matching actions. Perception is, of course, continuous, as is the lagged evaluation of half a second and the input that it generates in turn.

The combination of emotional weights constitutes an emotional mood, the outcome of which triggers behaviour on the continuum between aggression and flight on the one side, and submission on the other. With the experimentally observed shift of half a second, the outcome of this behaviour is displayed in the “aperture of consciousness”: these new percepts trigger new reminiscences, which in turn produce a new combination of emotional weights, leading to a new assessment open to action, and so on.

The body measures the power balance between itself and the prevailing conditions of the universe around it. Fear is an assessment with a negative balance; anger is an assessment with a positive result. The emotion felt reflects the difference between the hormonal mood that corresponds to the response deemed appropriate by the body and its actual ability to respond in the present moment.

Fear is therefore not, as a traditional conception would have it, an interference of emotion with motor behaviour; on the contrary, it is rather an adequate measure of the effort that will have to be expended in flight. When fear is gone, it means I'm ready. Similarly, anger is the measure of the immediate benefit

that the body can derive from a power balance that is, in this case, favourable to it.

The key to the role played by affect in motor behaviour is therefore the following: the present behaviour of the body is communicated to consciousness with a slight delay; this display itself - whose simultaneity of the percepts is guaranteed by the mechanism of *retrospective assignment* - triggers an evaluation of the body/world power balance, leading to corrections such as “flight forwards” or “flight backwards” (the meaning of these corrections is experienced as emotion, on a continuum that extends from fear to anger): as in speech, the landscape of the configuration space of the dual complementary system that we constitute with the world around us changes even as the process is unfolding. The actions taken therefore lead to a redefinition of what will happen next, thus ensuring effective control.

It is this banal observation that an effective control takes place, which reinforces us in the spontaneous belief (and endorsed by culture) that consciousness plays a decision-making role based on the goals it assigns itself. Consciousness plays indeed a role in the decision-making process, but as a mere cog in the display that regulates the emotional dynamics. It is the latter which actually ensures the entirety of decision-making, making it an exclusively unconscious process.

The discourse that the conscience holds itself, the *inner speech*, has no direct impact on our behaviour: it carries no true *intentionality* that would get materialised in our actions. *Intentionality* is an artefact generated by ex-post rationalisation, which could be viewed as a “superficial” version: only *worries* exist. However, the indirect impact of our inner speech is not negligible since it acts on our own emotional dynamics. As for our “outer” speech, it influences our interlocutors' emotional dynamics - regardless of the degree of our misunderstanding about our own motives which our speech automatically conveys: our interlocutors will always find there something to gauge the power balance prevailing between us - not at the level of conscious interpretation, but at the level where our respective emotional dynamics dialogue between them “peer to peer” and unbeknownst to our consciousness.

Consequently, it is only retrospectively that we discover in our own past behaviour, and by a rationalisation to the “second degree” this time, intermediate goals that we would have aimed at. In truth, it has only ever existed one step at a time: a progression that follows a descent gradient in a landscape of potentialities fulfilling a *concern* rather than achieving a *goal*.

* * *

Let us take the example of what is called “*eduction*.” “*Eduction*” is supposed to be a kind of “dress rehearsal” that the body can perform before a “voluntary” physical act takes place. Suppose I have to cross a ditch. Two cases are possible: I jump “without having to concentrate,” or I jump “having to

concentrate.”

If I jump without having to concentrate, my body jumps and informs my imagination of the jump that has just taken place. By retrospective assignment the jump and the awareness that I have of it appears simultaneous to my imagination and I have no difficulty in reconstructing what happened as a “decision to jump” followed by a “jump.”

If, on the contrary, I jump having had to concentrate, there is *eduction*. *Eduction* is lived by consciousness as a kind of mental “rehearsal”: I concentrate, and my body “mentally rehearses” the act to be performed until it is “ready.” Once I’m ready, I jump.

What does this mean in the new perspective opened by Libet's experiments and supplemented by our theoretical reflection? Once again, in this second case, as in the previous one, it is necessary that at the end of the day the explanation involves my jumping body, and then my imagination, which is informed of it.

I'm on the edge of the ditch, “about to jump,” but I'm not jumping yet. My body is “about to jump”; the signals announcing the jump come to my imagination as “mental rehearsal.” Which reminisces as a whole a bunch of memories associated with jumping in similar situations. For example, I miss my jump and bruise my face at the bottom of the ditch, and these evocations are enough to paralyse my body for a while. The level of adrenaline increases in my blood and changes my emotional mood, my “self-confidence,” until the memories evoked on the stage of my imagination are now those of the successful jump, and the triumph that accompanies it as its reward.

The paralysis of the body by reminiscence comes to an end and the body jumps. The imagination is informed of the act performed by the body: by *retrospective assignment* the course of events is construed as “I concentrated, and when I was ready, I jumped.”

If the act to be performed is difficult, it is possible that the imagination may be a hurdle by itself by continuing to associate signals from the body with a representation of failure. What we need to achieve in this case is, as any athlete knows, to “empty one’s mind”; in other words, that nothing takes place on the virtual stage of the imagination.

How does one practically “empty one’s mind”?

One concentrates... one stops thinking and one... empties one's mind...

Of course not, the explanation is inconsistent: the old ways of expressing oneself are difficult to get rid of because language as a whole participates in this spontaneous representation which stages consciousness as a decision-making body. As we will see for *attention*, “focusing” is most probably an accidental quality rather than a process. Quite simply, that the stage of the imagination where remembrances are called upon be deserted.

Common experience reminds us of that: if I suddenly “think” about the fact that I am running down the stairs at a brisk pace - that is, if I make of it a representation in my imagination - I am almost certain to break my neck, the emotion associated with the representation interfering with the act in progress.

In other words, *eduction* is in no way a “mental rehearsal” operated by consciousness, it is simply instead consciousness being informed of the body's hesitations. Fear can cause paralysis: if the tendency to flee exactly offsets the tendency to face danger, a transient equilibrium is established in inaction. Emotion accompanies remembrance, which is summoned in the imagination by the experience of the perilous situation. Vertigo, one hears sometimes say, is “the temptation to jump into the void.” It is more likely to be the paralysis that grips the body when the emotion, generated by an *imaginary* representation of the fall, inhibits the healthier reaction of the body to speed up to move away from danger. The fact that we think of ourselves as falling may have led us to mistakenly believe that we are tempted to leap. The sweat which then oozes in the palms of the hands is the beneficial preparation of the body for the eventuality of clinging to any object which will prevent, or at least interrupt, the fall.

* * *

There are therefore two types of decision-making processes for motor acts: those that are traditionally referred to as “voluntary” and those that are called “involuntary.” These do not give rise to “awareness,” of them consciousness is not informed: most of the time the beat of my heart passes unnoticed to me, it is only when the heart is racing fast, inducing vibrations that other receptors will register, that I become aware of these beats. The opposite is true for “voluntary” acts, which consciousness is informed of ex-post, but with a *retrospective assignment* which decoys consciousness and leads it to conceive the triggering of the act and the awareness of it as being synchronous and to imagine consciousness as being the source of the intention that was realised in the act.

In the reflex act, for example, the hand has been removed hastily after contact with the burning object - consciousness here records the movement of the body as having taken place, but without maintaining the illusion of an intention that would have presided over the act. The reflex act is being displayed in the *aperture* of the present moment, subjectively lived, only in exteriority; consciousness is informed of it ex-post but in a purely passive way, without assuming responsibility for its realisation. And this, because the decision-making took place here at a very early stage, preventing a coincidence in time - “construed” by retrospective assignment - of the act and its awareness, thus depriving consciousness of the illusion of intentionality. In a case like that of burning one's hand there is no benefit for the body in mobilizing the *experience* recorded in memory and accessible through reminiscence: burning one's skin is harmful, that, he knows, and without having to learn it.

As I already pointed out in the above passage, the naming of “voluntary” acts appears as usurped: if the word “will” designates the power of consciousness to be at the origin of motor acts, then Libet's work now establishes that the will is illusory, consciousness being deprived of such a power. No act should henceforth be characterised as either “voluntary” or “involuntary,” since it turns out that no act whatsoever results from the exercise of the *will*. Only the distinction between the two types of acts remains relevant, formerly called “voluntary” and formerly “involuntary,” but the difference between them boils down to this: that the former maintain in a subject the illusion of an authority such as the *will*, i.e. encouraging consciousness to delude itself as a decision-making power, while the latter, passing unnoticed by consciousness, are considered *a priori* as outside its jurisdiction.

Let us also note that when the “voluntary” motor act has become an automatism, consciousness ceases to be always informed of its fulfilment. “Where did I leave my keys?” “Did I turn the light off before leaving the house?” I have got no clue: putting my keys in my pocket or turning off the light are automatisms, i. e. “voluntary” acts that only rarely appear within the aperture of consciousness. *Attention* is then the name of a quality supposed to the subject when (s)he performs a “voluntary” act that gets displayed within consciousness. One says, “I changed the tire, and then I drove very carefully”: whenever this “voluntary” act is an automatism and its realisation is not displayed within consciousness, *attention* is conversely supposed to be lacking: one says then, “When the tire burst, we were chatting, I wasn't paying any attention to the road.”

As in the case of the reflex act mentioned above, the “voluntary” motor act, which has become an automatism, no longer needs an evaluation of the proper hormonal response: the body knows what to do. As a result, behaviour takes place under better conditions if imagination is missing: the “voluntary” act that has become automatism no longer benefits from reminiscence. Any modification of the hormonal equilibrium as feedback from the imagination is here parasitic: the imaginary representation would simply interfere with the act taking place, just as with the above-mentioned case of running down precipitously a flight of stairs.

No more than the *will*, therefore, *attention* is a faculty, its presence or absence is no more than *phenomenological*: simply signalling the display within consciousness of the type of act likely - in certain circumstances - of such a display. What is particularly lacking when “attention is absent” is the ability to rationalise, to hold a discourse of the *inner speech* type as a commentary on on-going events; as a rule, the unconscious, the *body*, remains perfectly aware of the location of the keys which are only “lost” as far as consciousness is concerned and retrieves them without difficulty as soon as “one ceases to think about it,” as soon as the automatisms recover their grip.

These remarks do not, of course, prohibit further use of words such as

intention, attention, concentration or *will* in their phenomenological use. There will always be people who are “capable of attention” and others who “demonstrate willpower”; there will always be acts of “deliberate intention” even if *attention, will, concentration, or intention* as mental faculties or dispositions do not exist. What is referred to by these expressions is the quality of some behaviour, in terms of its continuity, coherence, lack of self-contradiction, etc. However, our reflex of reading in these qualities the outcome of the exercise of a faculty must be dispensed with.

* * *

The unique language of human beings allows them to ensure that their needs and drives are met through speech alone: by asking someone. *Before* language exists, the animal must deal with it directly: by means of scenarios based probably entirely on associations of percepts and emotions, and anchored before any possible learning in these few basic “scripts” that we usually call *instincts*. As for us human beings, we no longer know what it was like to feel a need at a time when there was not yet the inner speech that saying “Here I am hungry!”

The prevalence of concepts in language, as well as the absence for the most part of a “significate” having an existence in the sensitive world, implies that the signified of one signifier is generally another signifier rather than a *significate*. Does this mean that the advent of human language has reduced the scope of sensory organ perception for survival? Does the fact that concepts can be linked to concepts restrict the role of external percepts of origin to be solely - according to Lacan's fortunate expression - “pad points”? Yes, to a certain extent, if only because some needs and drives are met precisely by uttered sentences (“I love you,” for example). For the rest, it does seem that, as opposed to the imagination, the body can do without words, and that for that reason they experience somewhat separate lives.

What consciousness perceives of the world, the world outside the skin of the speaking subject, is above all “what the words say about it,” the rest being borne by the body, which informs us of what it does only occasionally. Imagination now feeds on a material that is essentially linked to the storage of words in memory: as a result, consciousness “sees” through the grid of the concept; the body, meanwhile, busy with another matter - the mammal that we are - keeps us alive. It's the body, for example, which suddenly steers the wheel to avoid the coyote crossing the road, while I chatter, or rather - because that's what it's all about, and in any case - while “I'm listening to myself talking.”

What speech allows is the staging in the imagination of a universe absolutely cut off from all immediate physical resonance: the possibility for any speaker and at any time to produce by words the representation of a fictional universe. If I improvise a fairy tale for a child about to fall asleep, it is a reaction to a physical environment: there is indeed a causal relationship between the child

and the production of my story. But this physical relationship is confined to the contextual relationship between the mobilised *narrative genre* and the cultural situation of the child at bedtime. There is in no way any correspondence between this situation and the very *narrative content*: no correspondence between me sitting on the edge of the little bed and the dragons and enchanters who populate my story. Consequently, the content of the narrative is deprived of any *deixis*, of any “monstration,” it is *imaginary* in the sense that it takes place entirely in the imagination, and in no way reflects the physical circumstances of the narrator.

In other words, language allows the imagination to play entirely on the fictitious mode, disconnected from the immediate physical universe - while the body, in the meantime, quietly enables as they say, “maintains the ship.” In another context, the fact that a machine, namely a computer, can do without being completely in touch with the physical world that surrounds it, and nevertheless produce through software an adequate manipulation of sequences of words - sufficient in any case to achieve its specific functionalities - reinforces the hypothesis that it could be the same for human beings: that is to say, that language allows the bodily behaviour and verbal behaviour to develop in parallel, and may even be perfectly orthogonal: entirely dissociated from each other.

The intellectual often embodies such disconnection in at least a partial form. The image of the “absent-minded professor” offers a caricature of it in the popular representation: his body does one thing while his inner speech evokes a totally different one; the interferences provoked in his approach by the imagination and his own associations sometimes lead him to bump into a lamppost (scholars will add, “reacting with: ‘Pardon, Madame’”).

The same divide between spatio-temporal and linguistic behaviour actually exists at present between the distinct domains of robotics and Artificial Intelligence. A bridge between the two, which would allow an authentic simulation of human behaviour in its entirety, would require a dual system operating according to the mechanism revealed here. Namely: 1° the equivalent of consciousness identified as an *aperture* where percepts are displayed in real time, 2° the equivalent of a consciousness generating a discourse on what it perceives of the world and the behaviour of the body to which it is attached [this discourse being a *rationalisation*, i.e. an ex-post comment and deprived of any access to the true motivations of the body to which it is attached]. 3° the equivalent of a consciousness generating instructions constituting a disposition to action from the emotional weights associated with the words that make up the discourse it produces [the meaning of the latter, and in particular all that it implies as regards a decision-making power of the said consciousness being without consequence, consciousness constituting in this respect a dead end]. In other words - in more technical terms - the robot's actions would have to be displayed in an *aperture*, generate linguistic-type associations from this stage,

and produce a combination of signals that, when a threshold is reached, trigger an action, leading to the temporary relaxation of the system.

* * *

I wrote in “Jean Pouillon and the mystery of the Chinese room,” that “The parable of the ‘Chinese room’ backs the project of an Artificial Intelligence by staging something – be it man or machine - who, although not understanding the Chinese language, would *speak* it as well - or even “more correctly,” Pouillon would add - as somebody who would perfectly master its meaning” (Jorion 1997b: 95). Pouillon affirmed, in the passage to which I dedicated my remarks, “I do not know the meaning of the words they use... This ignorance is by no means a handicap...[it allows me] to forge sentences, often more ‘correct’ than those used by the speakers and whose grammatical structure guarantees that they can have a meaning. Which one? I don't know...”(Pouillon 1993: 155-157).

What is lacking, in the example presented by Pouillon, of a sense that has been maintained despite the meaning of the words invoked being void in his eyes, is the *imagination*, which is essential - as we have seen - to constitute in a speaking subject an emotional dynamics that feeds itself in speech, be it inner as well as outer. But the imagination is irrelevant when - as here - it is simply a matter of reporting someone else's words, i. e. acting as a simple screen whereupon the emotional dynamics of other people are projected. In this case, Pouillon's *imagination* would produce interference here, importing it as a parasitic effect, emotions linked to events of his own personal history to which “coils,” “long profiles,” “quarto plates,” are foreign, whereas what solely matters is the faithful *rendering* - whatever meaning anyone will assign the sentence.

Let us return then to these three theses that I thought it was possible to infer deductively from the “Pleasure of not understanding” and incidental remarks on related questions in his *Temps et roman*, theses about which I stated that they “display the obvious signs of absurdity” (Jorion 1997b: 96).

1° knowledge of syntax rules is essential for the correct use of a language,

2° knowledge of semantic rules is indifferent to the correct use of this language (ibid).

Indeed, *syntax* fails to reach imagination: it is of the same nature as the formerly called “involuntary” act; *semantics* is on the contrary what succeeds therein: it is of the same nature as the formerly called “voluntary” act. It is this contrast that had allowed some Scholastics, especially Jean Buridan, to assert that syntax is deprived of meaning: it has a meaning but which is not what we call *signification*, which is precisely what semantics exclusively conveys. The meaning of syntax is entirely dealt with upstream of consciousness, at the unconscious level, by the body - according to the expression I have used in this text -, it is the blind, inaccessible part of meaning, which forces us to contort

our minds when we try to define - to use the scholastic vocabulary - a *syncategorem* such as “nonetheless.”

3° The intuitive sense of understanding is indifferent to the correct use of a language “(ibid.).

Understanding is the evocation by *imagination*, activating at the same time as the concepts evoked, all those connected with them (cf. Jorion 1990a: chapter 9). The content of this representation is *signification*, but there is a part of the meaning that is absent from *signification*: all that is of the order of the structure, of the frame of the sentence; the structure does not pass through the evocation, that is what Freud observes about the dream: that syntactic effects must be expressed as a rebus, that they must be evoked indirectly, in a figured, figurative form. In other words, one is forced to reproduce the meaning of *syncategorems* by means of *categorems* assembly.

This proves that the production of “correct” sentences can actually take place without anything appearing on the stage of the imagination of the person who generates them. Hence the reality of the effect observed by Pouillon: he produces “sentences that are often more correct” by entrusting the task to his *body* and letting his *imagination* wander around on conceptual objects that are probably “more exciting” than “coils,” “long profiles,” “quarto plates,” himself, remaining as though “outside of communication.”

* * *

At the time I suggested a “rectification of names” where the word “unconscious” is replaced by the word “body” and the word “consciousness” replaced by the word “imagination,” some readers may have taken the next step of replacing the word “imagination” with the word “soul.” This step is tiny and can now be taken without major objection. We would then have gone from the concepts of psychoanalysis in terms of “conscious” and “unconscious” to the much older “body” and “soul” dichotomy, corresponding to what I called a spontaneous psychology written in our language.

The only difference, but it is essential, is that here we are dealing with an inverted relationship compared to the classical scheme in terms of responsibilities and decision-making: here, the soul is not in control, it is the body; the soul is entirely subordinate to it.

What the parable of the Chinese chamber shows is a prisoner who speaks Chinese, having made the Chinese semantics – against his will - a simple element of the syntax of the language. His body speaks Chinese, and his soul is in no way informed of it. What does the soul do, meanwhile? It's hard to tell. However, if I were this soul, I would dream of escaping from the jail. I might be humming the refrain of a ballad by Johnny Cash with Sartrian accents: “The walls of a prison will never hold me...” That said, if the prisoner of the Chinese chamber ever managed to escape, it would be his body that would not only

have succeeded, but, as we now know, would actually have taken the initiative and performed every necessary gesture.

This conclusion we reached did not require - as one may have noticed - the experimental verification whereof Benjamin Libet is the author (which does not take anything away from his immense merit): it was possible to achieve this in a deductive way from a reflection on inner speech, and after replacing one logic in terms of *final causes* with another in terms of a *gradient*.

* * *

Why did two thousand five hundred years of reflection prove to be powerless to question the decision-making power of consciousness? It seems to me that there is, on this issue, something of the order of prejudice, of what changes only as a last resort in conceptual organisation (what I have called elsewhere a *core belief*), something of the order of a *taboo*.

The history of science may be of some help here. When Max Planck lays the foundations for quantum mechanics, we can easily list his predecessors: he builds on the foundations laid by Clausius, Maxwell and Boltzmann. When Darwin develops his theory of species evolution, or when Freud develops Freudian metapsychology, it would conversely be difficult to determine their predecessors (the parallel work of Wallace is contemporary with Darwin's). However, they can be found here and there in history (and sometimes only a few years before), *precursors*, thinkers who expressed views where one finds in germ, in sketched form and most often in isolated ideas, what will take all its meaning only in the complete theory that Darwin or Freud then developed. When *predecessors* exist, as in the case of Planck, the search for *precursors* would appear to be futile, since a crowded continuous line of predecessors would lead to them.

What distinguishes then Darwin's or Freud's discoveries, if not their real novelty? "That they do not constitute theories as such," say some people today, "because they are not falsifiable and do not lend themselves to counter-proofing." The argument is without merit: their theories are falsifiable, like the Big Bang for example, even if it would require more discursive argumentation than pure experimental verification.

What distinguishes Darwin's or Freud's discoveries, if not their real novelty?" That they do not constitute theories as such," say some people today," because they are not falsifiable and do not lend themselves to counter-proofing." The argument is without merit: their theories are falsifiable, like the Big Bang for example, even if it would require more discursive argumentation than pure experimental verification. What distinguishes their constructions is that it is difficult, in the sense of "tough" psychologically, for an author to formulate them. There is here, as I said, a *taboo* to overcome, something that provokes fear or anger if touched upon: there is also a *conversion* to take place, firstly for its author, at the moment when he formulates his theory, and

secondly for his reader at the moment when he must be convinced, when certain ramparts erected by his emotional framework need to collapse in order to make room for the new conception.

What characterises Darwinism or Freudianism is that if they are true, the merit of Darwin and Freud, in the capacity of being their author, is automatically diminished. If we are only the descendants of great apes, then Darwinism itself has as its author the descendant of a great ape (the cartoonists of the time joyfully indulged along that line, by the way), likewise, if every human work is a diverted means of satisfying a sexual impulse, then Freudian metapsychology itself is a diverted means for its author to satisfy such an impulse.

Darwin's theory of evolution, and psychoanalysis as well - as it has been written - imply a downgrading, a *belittling* of the human race's self-image. The vanity of the species takes a nasty blow, for it is much more than a new theory; it is also a lesson in humility. Copernicus had done the same when he moved the Earth from the centre to the periphery or when Linnaeus first classified man as ranking with mammals.

So what was preventing us from understanding the real distribution of responsibilities between body and soul? Probably a psychological mechanism similar to the one I just mentioned about Darwin and Freud: if this is the case, then composing the Ninth Symphony or painting *The Night Watch*, are undoubtedly *personal* achievements with their foundation in a biological being shaped by a history, but which are no more related to a human *subject* master of his actions, than the fact for anyone among us to open a window *mechanically*. As for the one who would attach his name to the discovery that the functions of the soul and the body must simply be reversed, he would *belittle* this way his own discovery: it would have been just as mechanical, according to the automatism that he would have highlighted. He would be the author of his discovery by a mechanism whose - he would have proved it - his own person is the medium only for reasons that are perfectly incidental in the light of history. Anything he could say about the fact that it could only take place through him would be automatically disqualified: it could likely be only him, but without the discovery's fatherhood going back to that "me, myself, I" with which he likes to punctuate his speech.

This, in a nutshell, would explain why the thinkers who looked into the mystery of the Chinese chamber stopped at the brink of its elucidation - since what was to be discovered would have deprived them of the satisfaction of putting forward their own person - a satisfaction that has always guided the process of discovery.

Laguna Beach, 3 December 1997

References

- Chalmers, David J. *The Conscious Mind. In Search of a Fundamental Theory*, Oxford: Oxford University Press, 1996
- Churchland, Patricia Smith, On the alleged backwards referral of experiences and its relevance to the mind-body problem. *Philosophy of Science*, 48, 1981: 165-181
- Dennett, Daniel C., *Consciousness Explained*, Boston: Little, Brown and Company, 1991
- Derrida, Jacques, *La dissémination*, Paris: Minuit, 1972
- Dumont, J.-P., *Les écoles présocratiques*, Folio, Paris: Gallimard, 1991
- Freud, Sigmund, *Introduction à la psychanalyse*, Paris: Payot, 1949 [1917]
- Octave Hamelin, *Le système d'Aristote*, Paris: Vrin, 1985 [1904-1905]
- Jorion, Paul, « Intelligence artificielle et mentalité primitive. Actualité de quelques concepts lévy-bruhliens , » *Revue Philosophique*, 4, 1989: 515-541
- Jorion, Paul, *Principes des systèmes intelligents*, Collection Sciences Cognitives, Paris: Masson, 1990a
- Jorion, Paul, « An alternative neural network representation for conceptual knowledge , » communication présentée à la British TELECOM, CONNEX Conference, Martlesham Heath, January 1990b, 23 pp; <http://cogprints.soton.ac.uk/abs/comp/199806036>
- Jorion, Paul, « L'intelligence artificielle : au confluent des neurosciences et de l'informatique , » *Lekton*, vol. IV, 2, 1994a: 85-114
- Jorion, Paul, review of Jean Pouillon, *Le cru et le su*, *L'Homme*, 130, 1994b: 150
- Jorion, Paul, « La linguistique d'Aristote , » in V. Rialle & D. Fisette (eds.), *Penser l'esprit : Des sciences de la cognition à une philosophie cognitive*, Grenoble: Presses Universitaires de Grenoble, 1996: 261-287
- Jorion, Paul, « Ce qui fait encore cruellement défaut à l'Intelligence artificielle », *Informations In Cognito*, No 7, 1997a: 1-4
- Jorion, Paul, « Jean Pouillon et le mystère de la chambre chinoise », *L'Homme*, 143, 1997b: 91-99
- Jorion, Paul, « Le moment du Verbe : Le signifiant et son efficace », *L'Homme*, 145, 1998: 239-248
- Jung, Carl, « Psychoanalysis and association experiments », 1906, in Carl Jung, *The Collected Works, Volume Two: Experimental Researches*, London: Routledge & Kegan Paul, 1973: 288-317
- Kneale, W. & Kneale, M., *The Development of Logic*, Oxford: Clarendon Press, 1986 [1962]
- LeDoux, Joseph, *The Emotional Brain. The Mysterious Underpinnings of Emotional Life*, New York: Simon & Schuster, 1996
- Lévy-Bruhl, Lucien, *La mythologie primitive*, Paris: PUF, 1963 [1935]
- Libet, Benjamin, « The experimental evidence for subjective referral of a sensory experience backwards in time: Reply to P. S. Churchland », *Philosophy of*

- Science, 48, 1981: 182-197
- Libet, Benjamin, « The Neural Time-Factor in Perception, Volition and Free Will », *Revue de Métaphysique et de Morale*, 2, 1992: 255-272
- Libet, Benjamin, « Conscious Mind as a Force Field: A Reply to Lindahl & Århem », *Journal of Theoretical Biology*, 185, 1997: 137-138
- Merleau-Ponty, Maurice, « Sur la phénoménologie du langage » (1951) in *Signes*, Paris: Gallimard, 1960: 105-122
- Moody, E. A., *Truth and consequence in Mediaeval Logic*, Amsterdam: North-Holland, 1953
- Nietzsche, Friedrich, *Le crépuscule des idoles ou comment on philosophe au marteau*, Paris: Denoël / Gonthier, 1970 [1888]
- Pouillon, Jean, *Temps et roman*, nouvelle édition augmentée, Tel, Paris: Gallimard, 1992 [1946]
- Pouillon, Jean, *Le cru et le su*, Paris: Seuil, 1993.
- Roudinesco, Élisabeth et Michel Plon, *Dictionnaire de la psychanalyse*, Paris: Fayard, 1997
- Searle, John R., *Minds, Brains and Science*, The 1984 Reith Lectures, London: BBC, 1984
- Searle, John R., « Consciousness and the Philosophers », *The New York Review of Books*, Vol. XLIV, 4, March 6, 1997: 43-50
- Wittgenstein, Ludwig, *Zettel*, Oxford, Basil Blackwell, 1967