SEARLE'S CRITIQUE OF THE MULTIPLE DRAFTS MODEL OF CONSCIOUSNESS

UDC 81'23(049.32)

Dorde Vidanović
Faculty of Philosophy, University of Niš, Serbia
E-mail: vidanovic@gmail.com

Abstract. In this paper I try to show the limitations of John Searle's critique of Daniel Dennett's conception of consciousness based on the idea that the computational architecture of consciousness is patterned on the simple replicating units of information called memes. Searle claims that memes cannot substitute virtual genes as expounded by Dennett, saying that the spread of ideas and information is not driven by "blind forces" but has to be intentional. In this paper I try to refute his argumentation by a detailed account that tries to prove that intentionality need not be invoked in accounts of memes (and consciousness).

Key words: Searle, Dennett, Multiple Drafts Model, consciousness, memes, genes, intentionality

"No activity of mind is ever conscious"
(Karl Lashley, 1956)

1. INTRODUCTION


Received June 27, 2009
1 A version of this paper was submitted to the Department of Philosophy of the University of Maribor, Slovenia, as part of the Festschrift for Dunja Jutronic in 2008, see http://oddelki.ff.uni-nb.si/filozofija/files/Festschrift/Dunjas_festschrift/vidanovic.pdf
Searle claims that Dennett bases his explanation on four notions, von Neumann machines, connectionism, virtual machines and memes. Interestingly enough, Searle emphasizes that this AI view cannot be true mostly because of the misguided metaphor of memes standing for virtual genes. In this paper I plan to examine Searle's critique of this metaphor as expounded in Dennett's Multiple Drafts conception of consciousness, as argued by Dennett during the greater part of the nineties and, in another form, and under another name, in the last couple of years, excepting the Rohwolt published paper that came out in Germany in 1996 (cf. Dennett 1991, 1996, 1998, Dennett and Kinsbourne 1992; Dennett 1996, 2001, 2005).

2. THE MULTIPLE DRAFTS MODEL IN BRIEF

To clarify the notion of the so-called MDM (Multiple Drafts Model) of consciousness, one can turn to Brook and Ross who state that Dennett's vision of the conscious system is typically Cartesian, resembling a transparent screen "on which conscious states play before a little homunculus sitting in the middle of the theatre... where the conscious states themselves are conceived of as discrete, separately identifiable states, states with, for example, clear stop and start points. Dennett wants to replace the Cartesian picture with what he calls a Multiple Drafts Model (MDM) of consciousness. The MDM treats consciousness as a kind of mental content, almost a matter of programming, a highly controversial point of view." (Brook and Ross, 2002: 8).

It appears that the Multiple Draft Model (MDM) of consciousness was developed as an alternative account to the intuitively attractive, "but incoherent, model of conscious experience Dennett calls Cartesian materialism, the idea that after early unconscious processing occurs in various relatively peripheral brain structures "everything comes together" in some privileged central place in the brain which Dennett calls the Cartesian Theater -- for "presentation" to the inner self or homunculus" (Dennett & Akins 2008: 4321). Dennett opposes the postulation of such a locus, but, as he claims, there are theories that simply take its existence for granted.

He states that such 'consciousness-oriented' theories "postpone indefinitely the task of saying where and when the results of all the transformations and discriminations are 'made available to conscious awareness,' which suggests sending their products to some higher center (for what purpose? what would happen there?)" (Dennett & Akins 2008: 4321), or, else, such theories argue that "a decision to move a limb takes several hundred milliseconds to 'rise to consciousness,' creating an ominous picture of human agents as deluded about their ability to make a conscious decision" (Dennett & Akins 2008: 4321).

There is yet something else that "dualist" or, presumably, "interactionist" theories (theorists) presuppose and that is the idea that the transfer by human senses into an unconscious neural code must be followed by yet another conversion into some other "medium", which Dennett, not unexpectedly, dubs consciousness.

Thus we can see that the MDM opposes the existence of a singular locus of convergence claiming that there is an array of information states that occur simultaneously in the brain, just like multiple drafts of a scholarly paper. Dennett pursues the point succinctly, using the concept of the meme in his seminal work regarding consciousness:

"Human consciousness is itself a huge collection of memes (or more exactly, meme-effects in brains) that can best be understood as the operation of a "von Neumanesque"
virtual machine implemented in the parallel architecture of a brain that was not designed for any such activities" (Dennett 1991: 210).

3. SEARLE'S CRITIQUE

Leaving aside for the moment the fact that Dennett's descriptions of such a connectionist net or parallel distributed processing in the guise of neural nets do seem plausible, at least prima facie, I should like to focus on the notion of 'meme' that Dennett liberally uses and, at the same time, attempt to analyze it in the light of Searle's global critique of the MDM. Regarding the meme and Dennett's usage of the concept, Searle writes that "this notion is not very clear. It was invented by Richard Dawkins to have a cultural analog to the biological notion of a gene. The idea is that just as biological evolution occurs by way of genes, so cultural evolution occurs through the spread of memes." (Searle 1997: 104). As Searle goes on to quote Dawkins's definition of the meme as quoted by Dennett, I will, as a substitute, try to present a clearer description of this important concept by quoting Dawkins himself:

"The gene, the DNA molecule, happens to be the replicating entity that prevails on our own planet. There may be others. If there are, provided certain other conditions are met, they will almost inevitably tend to be become the basis for an evolutionary process.

But do we have to go to distant worlds to find other kinds of replicator and other, consequent, kinds of evolution? I think that a new kind of replicator has recently emerged on this very planet. It is staring us in the face. It is still in its infancy, still drifting clumsily about in its primeval soup, but already it is achieving evolutionary change at a rate that leaves the old gene panting far behind.

The new soup is the soup of human culture. We need a name for the new replicator, a noun that conveys the idea of a unit of cultural transmission, or a unit of imitation. 'Mimeme' comes from a suitable Greek root, but I want a monosyllable that sounds a bit like 'gene'. I hope my classicist friends will forgive me if I abbreviate mimeme to meme." (Dawkins 1989: 192).

Searle claims that the analogy between the gene and the meme is mistaken. He states that evolution advances by brute, blind and purely nature-driven forces and elements. He goes on to say that "[The] spread of ideas and theories by "imitation" is typically a conscious process (emphasis mine) directed toward a goal. It misses the point of Darwin's account of the origin of species to lump the two sorts of processes together. Darwin's greatest achievement was to show that the appearance of purpose, planning, teleology, and intentionality in the origin and development of human and animal species was entirely an illusion. The appearance could be explained by evolutionary processes that contained no such purposes at all. But the spread of ideas through imitation requires the whole apparatus of human consciousness and intentionality" (emphasis mine) (Searle 1997: 105).

Searle is adamant about the above-mentioned: "In short, the transmission of ideas through imitation is totally unlike the transmission of genes through reproduction, so the analogy between genes and memes is misleading from the start" (Searle 1997: 105).
4. A LIMITATION OF SEARLE’S CRITIQUE

There is a serious flaw in Searle’s objections directed at the supposed similarity between genes and memes granting that they are vehicles of imitation and/or reproduction. To rehash his argument: Searle believes that the spread of ideas, concepts and behavior through imitation (via memes) must be intentional thus requiring awareness and conscious perception on behalf of human subjects.

In spite of his firm convictions, here Searle finds himself on a shaky ground. In order to see where his objections to Dennett’s account of memes go wrong, let me analyze some of the issues related to the nature of the propagation of informational units such as memes described by Dennett in many of his papers and books.

Firstly, the original text that started this all, Dawkins’s *The Selfish Gene* never states in an explicit manner that the spread of memes is exclusively a conscious process. Dawkins keeps both options (conscious and unconscious proliferation of memes) open: "If the idea catches (sic!) on, it can be said to propagate itself, spread from brain to brain. As my colleague N. K. Humphrey neatly summed up an earlier draft of this chapter: ‘…memes should be regarded as living structures, not just metaphorically but technically. When you plant a fertile meme in my mind you literally parasitize my brain, turning it into a vehicle for the meme’s propagation in just the way that a virus may parasitize the genetic mechanism of a host cell’" (Dawkins 1989: 192). Thus Dawkins speaks of ideas catching on, their propagation, their being planted and their fecundity, ideas as parasites, etc., All of these terms indicate that Dawkins (and, by extension, Dennett) is ever so ready to acknowledge that any spread of ideas (memes) can be unconscious, and need not require intentionality as claimed by Searle.

Secondly, at the time Dawkins had *The Selfish Gene* published, he had no notion of the forthcoming neurocognitive research which would turn out to be compatible with his ‘indifferent’ approach to the issue of intentionality and compulsory awareness in humans presumably indispensable in the process of meme propagation.

For the past ten years there have been many research projects dealing with the problem of the nature of information processing. The issue has turned out to be much debated and controversial and the prevalent opinion favors ‘deliberation’ and ‘awareness’. The "novel idea" that we are not all the time aware of our internal cognitive processing was provocative at first but after the dust settled it became clear that information processing was not solely cognitive but had much to do with emotions and the unconscious. This broadened intellectual 'landscape' was the product of the research that gave birth to the idea that our cortex gathers and processes some indispensable emotional information that is all but deliberative. Consequently, the increased understanding of the neural underpinnings of some conscious and non-conscious processes has recently led to new insights into the workings of the human mind, most notably in the area of recognition of facial expression. In this field of research functional neuroimaging studies in both healthy subjects and brain-damaged patients come together and imply that the amygdala and the orbitofrontal cortical areas are activated by faces expressing emotion "independently of voluntary attention and even without any awareness of the stimuli presented" (de Gelder 2005: 124). Apparently, there is a primitive emotion system that operates independently

---

3 As a coverage of Freud and his account of the human mind is outside the scope of this paper, I have decided to leave it out at this time.
of deliberation, awareness and the higher states of the central nervous system: a kind of emotional alternative processing route or an evolutionary older neural shortcut.

One of the more provoking and controversial ideas here may be the hypothesis that the amygdala activity incited by emotional stimuli could be prompted by subcortical inputs independent of the preliminary analysis in the striate cortex and the frontal visual areas. "This 'low' route (as it is sometimes called) conveys rather crude information, based on a coarse parsing of a face stimulus; [and…] in itself, is sufficient to trigger an emotional response. If conscious and unconscious processes have their own neural networks […], then it is theoretically possible that both routes interact and that conflicts can arise between conscious and unconscious processes. Such conflicts, in turn, can be seen against the background of the larger question regarding the role nonconscious emotional processes play in daily life (emphasis mine) (de Gelder 2005: 125).

De Gelder's and Adolphs's research during the end of the nineties of the previous century and the beginning of the twenty first century, together with tens of present-day studies (for a sample see, e.g., Adolphs, Baron-Cohen & Tranel 2002) resemble the situation in which humans and primates with striate cortical damage exhibited a surprising degree of functional vision (previously researched by Humphrey and Weiskrantz at the beginnings of the 70's of the previous century).

What is at stake in both cases is the idea that our perception and cognitive processing are deliberate and (mostly) aware, with little or no unconscious or nonconscious emotional coloring and/or involvement of the higher central nervous system strata. The two research projects that I have just described shed substantial doubt on the deliberate and sentient functioning of the human mind and point in the direction of a significant emotional and non-deliberate portion of cognitive processing in the human cortex.4

Leaving the important emotional aspect aside for the moment, I am convinced that it is well-nigh impossible to think that our cortex might explicitly process (encode and decode) the overwhelming amount of information that we constantly phenomenologically experience at any given moment in time.

There is yet another area of prime importance with regard to the propagation of memes in which it is obvious that a very potent memetic set (natural language) appears to function mostly unconsciously and with little involvement from the conscious parts of the brain. An illustration is in order: during the process of language acquisition, especially in the so-called 'immersion' cases subjects simply 'pick up' a language and start using it as their own (evidenced by a plethora of real world cases in immigrant communities throughout the globe). In contemporary studies dealing with the process of foreign language acquisition and/or learning particular care is taken to make a precise distinction between the two as acquisition signifies an unconscious process while learning is specifically conscious and deliberate. It was early Chomsky and the supporters of the idea that the mind is not a blank slate (Chomsky 1957 and later, Jackendoff 1994, Pinker 1994, 2002) who emphasized the nature of spontaneous acquisition, ascribing the behavioral nature of the habit-formative process to conscious language learning.

A number of investigators have explored some of the threads arising from the analysis of a common pool of concepts. Frank C. Keil and his colleagues at Cornell University (Keil 1992, and a host of studies later) have inferred from the data collected that certain

---

4 Nothing new here, we can only reiterate Lashley's words from 1956 here. Or, else, Fodor's claim that "... practically all psychologically interesting cognitive states are unconscious..." (Fodor 1983: 86).
highly abstract concepts such as the ideas of *being alive* or *ownership*, most notably, carry a genetically inherited value in the young child's mind-set. Apparently, when the specific words for owning, giving and taking, keeping and hiding, and their relatives 'populate' the mind of a child, they find suitable niches (we might even dub them 'meme pools' as a counterpart of 'gene pools'). Ray Jackendoff and other conceptual semanticists have identified important structures of spatial representation possibly streamlined to enhance the control of *locomotion* and the *placement* of movable things. These are the cornerstone of our intuitions about abstract concepts such as *beside*, *on*, *behind*, and the like.

In addition to the research in cognitive processing, Nicholas Humphrey has recently claimed that there must be a genetic predisposition for the development of a consciousness that is *predominantly emotional* rather than deliberate and cognitive (Humphrey 1992), and Leslie, Knobe and Cohen (in press) have developed evidence for this, in the form of what they call a "theory of mind module" designed to generate second-order beliefs (beliefs about the beliefs and other mental states of others). These three authors also claim that "the theory-of-mind mechanism generating attributions of intentional action may have a parameter for the moral valence of outcomes (and perhaps for other kinds of valence). The value of this parameter would influence judgments of purpose, but still be obtained from processes external to a theory of mind, such as moral judgment" (Leslie, Knobe & Cohen (in press: 8). Some autistic children, for example, seem to be well-described as suffering from the disabling of such a module, for which they can occasionally make interesting compensatory adjustments.

To put another nail in Searle's concept of Dennett's usage of 'meme', let me mention that some researchers used the idea that the original memes (as exhibited among animals) can be defined as behaviors ready to be mimicked (e.g. Dawkins referred to bird songs as memes).

"[Nevertheless…] one reviewer remarked that only humans can imitate in an observable manner. If only this kind of conscious imitation counts for memes, then only humans produce memes and washing sweet potatoes by Japanese macaques would not be caused by imitation and thus not be memetic. One may object that there is strong evidence for unconscious imitation underlying learning in animals (emphasis mine), as becomes apparent from the work of Rizzolatti et al." 5:

"In area F5 of monkey premotor cortex there are neurons that discharge both when the monkey performs an action and when he observes similar actions made by another monkey or by the experimenter. We report here some of the properties of these 'mirror' neurons and we propose that their activity 'represents' the observed action. We posit, then, that this motor representation is at the basis of the understanding of motor events." (Rizzolatti et al. (1996: 135)

A rather speculative and tentative suggestion states that "conscious imitation itself might be a secondary consequence of the development of language, which makes possible reflexive awareness. If one could show that conscious imitation is a consequence of reflexive awareness (i.e., consciousness), this kind of imitation could be considered itself largely as explained once one has explained language." (Vaneechoute and Skoyles 1998).

Further research by Rizzolatti and Arbib (1998) and Arbib (2005) concerns the neural and functional grounding of the human language skill and its emergence in hominid evolu-

---

5 Vaneechoute and Skoyles 1998 (web paper from Journal of Memetics).
It transpires that both the premotor area F5 in monkeys and the "corresponding" area in humans (Broca's area) own a 'mirror system' employed in the execution and observation of manual actions. The researchers concluded that the F5 area in monkeys has its homologous counterpart in Broca's area, but the development never continued in the evolutionary direction of Homo sapiens. Still, the inability of primates to distinguish between acts of 'doing' and 'perceiving', as well as their inability to consciously detect their own neural activity in the primary visual cortex, as already observed by Humphrey and Weiskrantz three decades ago (Humphrey and Weiskrantz 1967), indicate that both our biological 'ancestors' and contemporary humans need not be aware, conscious or capable of introspection.

5. CONCLUSION

It goes without saying that conscious imitation itself might be a derivative consequence of the development of language, the one that makes reflexive awareness possible. If one could show that conscious imitation is a consequence of reflexive awareness (in other words consciousness itself), this kind of imitation may be considered as explained once we managed to understand the notion of language. Language is, as shown above, possibly a product of unconscious imitation, existing within a niche and hosting a gamut of emotional, moral and social habits, and, last, but not least, depending on the functioning of the biological parameters.

"[...one] may usefully think of words -- the most effective vehicles for memes-- as invading or parasitizing a brain, not simply being acquired by a brain. What is the shape of this environment when words first enter it? It is definitely not an even playing field or a tabula rasa. Our newfound words must anchor themselves on the hills and valleys of a landscape of considerable complexity. Thanks to earlier evolutionary pressures, our innate quality spaces are species-specific, narcissistic, and even idiosyncratic from individual to individual." (Dennett 1994).

I believe that Searle got at least one part of the memetic story right: when we talk about information units that are vehicles of imitation, then we primarily think in terms of words, as Dennett noticed above. Searle is right that we are most probably aware of at least a portion of ideas camouflaged as words, but there is also another, unconscious, part of the input to our brains that we take for granted due to many reasons, be they our emotions, our moral judgments or cultural heritage. Mimicking and imitation, in the widest possible sense, are most frequently carried out by way of language, while other modes of transfer can be attributed to other semiotic vehicles, such as fashion, hair styles or socially induced fads and demonstrations of crude political and financial interests.

To what extent language proves to be a deliberate, conscious and transparent mode of action is impossible to say with any degree of precision, but, to be on the safe side yet again, let us suppose that the conscious / unconscious (non-conscious) division of information transfer is unevenly distributed in humans. If this is true, as the previous argumentation seems to substantiate, then Searle's idea that there is no similarity between genes and memes due to the former being driven by the blind and incomprehensible forces of nature, and the latter operating in a transparent, non-opaque and deliberate fashion, is simply wrong.

6 These ideas reflect Dennett's analyses outlined in his seminal book Consciousness Explained (1991), particularly in chapters 7 and 8. The quotation is from his essay "The Role of Language in Intelligence", in Khalfa 1994.
John Searle seems to have assumed his proverbially vigorous 'common sense' approach in criticizing Dennett's description of consciousness that was partly accounted for by postulating memes as basic symbolic or information units (of imitation) analogous to genes (Dawkins 1989). Although common sense has helped Searle very often in the past, it is fair to say that this time common sense and facts have gone their separate ways. This is so because even though common sense dictates that thoughts and awareness/deliberation go hand in hand, this has been proven wrong in the past two decades of cognitive research, starting with Jackendoff's Intermediate-Level Theory of Consciousness (Jackendoff 1987), and continuing with recent Crick and Koch's work on the neuronal correlates of consciousness (Crick and Koch 2000, 2003).

Briefly told, Jackendoff's story offers a startling account of the inner workings of human consciousness, claiming that we are not aware (conscious) of our thoughts, but still manage to get an insight by the help of their representation in sensory terms. To achieve this feat, Jackendoff had to resort to a three-partite structure of mind/brain, postulating a third entity, called 'the phenomenological mind' which is responsible for the persistent illusion that we are directly aware of both reality and our inner world of thoughts. "Jackendoff remarks that common sense seems to tell us that awareness and thought are inseparable and that introspection can reveal the contents of the mind. He argues at length that both these beliefs are untrue. They contrast strongly with his conclusion that thinking is largely unconscious. What is conscious about thoughts is visual or other images, or talking to oneself. He maintains that visual and verbal images are associated with intermediate-level sensory representations, which are in turn generated from thoughts by the fast processing mechanisms in short-term memory. Both the process of thought and its content are not directly accessible to awareness." (Crick and Koch 2000: 106).

An example from linguistics may clarify this position. A bilingual person can express a thought in either of the two languages, but the thought itself, which produces the verbal activity or imagery, is not at her/his disposal for direct perception, although it can be accessed in the corresponding language guises (sensory forms).

As for the neuronal correlates of the human mind and the 'unconscious homunculus' perching up there somewhere, Crick and Koch state that humans are not directly aware of the outer flow of sensory events (or the 'world' as we experience it). Rather, they are conscious of the results of computations performed by the central nervous system on the neuronal representations of the sensory world (just as in the bilingual example above). Following in Jackendoff's footsteps, they claim that: "[...] these results are expressed in various cortical areas. Nor are we directly aware of our inner world of thoughts, intentions and planning (that is, of our unconscious homunculus) but -- and this is the surprising part -- only of the sensory representations associated with these mental activities. What remains is the sobering realization that our subjective world of qualia --- what distinguishes us from zombies and fills our life with color, music, smells, and other vivid sensations --- is probably caused by the activity of a small fraction of all the neurons in the brain, located strategically between the outer and the inner worlds" (Crick and Koch 2000: 110).

In their more recent paper dealing with the neural correlates of consciousness (NCC), they propose the so-called 'snapshot hypothesis', claiming that the conscious perception of motion is not represented by the change of the firing rate of the neurons responsible for the perception of the act, but by the frequent, almost incessant firing of the neurons that

---

7 With the exception of the primary visual cortex, cf. Crick and Koch 1995.
represent the motion (Crick and Koch 2003: 122). This may only mean that a figure in 'virtual motion' (such as a drawing on paper) will be perceived as being really in motion, although it is static in the 'real world'. This discrepancy of the NCC and states of the real world and/or sensory forms in which we experience reality sheds more light on the feature of replicating information units that is so easily overlooked in folk psychology: their tendency to create wholly new units and thus replicate, without any grounds in the workings of the 'real world'. The corollary is, naturally, the fact that such units tend to be unconscious, reflecting our 'phenomenological mind' as the mediator between the neuronal computation and the sensory forms that we take to be objects of our (mediated) awareness.

Folk psychology can also be linked to memes and, perhaps, be the target of philological and anthropological research attempting to clarify the nature of several well-known and much discussed aspects of oral traditions: traditional referentiality, anaphora, and the use of repeated metrical patterns. It goes without saying that all three phenomena, different as they are, can be understood as arising from the operations of the unconscious human mind, embedded in diverse ethno-anthropological associations (Drout 2006).

In closing, I hope to have shed some light on one particular, albeit minuscule, thread of Searle's critique of Dennett's Multiple Drafts Model of consciousness, showing that it is based on a misunderstanding of the nature of the gene-meme analogy as well as on a misinterpretation of the tenets of folk psychology that are always 'memetically' articulated ("thoughts are always transparent to the one who harbors thought; "all mental activities are conscious", "humans are rational", etc.).

REFERENCES

* And, per extensionem, the newly formulated 'fame in the brain' modification of the same, drawing on the same idea (of gene and meme replicators) started in Dawkins (1989), formulated in Dennett, D.C. (2005).

O OGRANIČENJU SERLOVE KRITIKE DENETOVOG VIŠESLOJNOG MODELA SVESTI

Dorde Vidanović

U ovom radu pokušavam da pokažem koja su ograničenja Serlove kritike Denetove koncepcije svesti zasnovane na zamisli da je komputaciona arhitektonika svesti modelovana poput proste replikacije jedinica informacije koje Denet naziva "memama". Serl tvrdi da meme ne mogu zameniti postojanje virtualnih gena kako to zagovara Denet, i isto če da su ideje i informacije uvek intencionalne, a ne neakvne "slepe sile" koje njima upravljaju. U radu pokušavam da takve argumente detaljno pobijem i da dokazem da nije nužno koristiti intencionalnost u objašnjenju informacionih replikatora.

Ključne reči: Serl, Denet, višeslojni model svesti, svest, meme, geni, intencionalnost