

JERRY FODOR'S PHILOSOPHY OF MIND

La filosofía de la mente de Jerry Fodor

LEOPOLDO EDGARDO TILLERÍA AQUEVEQUE*

INACAP, Temuco, Chile

leopoldo.tilleria@inacapmail.cl

Orcid code: <https://orcid.org/0000-0001-5630-7552>

Abstract

The article presents an approach to the philosophy of mind of Jerry Fodor, and focuses on the problem that his thesis on computational naturalism represents. It emphasizes the ontological differences between the modular input systems and the central systems of this mental machine. Through a review of the main criticisms of the Fodorian program, it is concluded with the idea that its conceptual innateness turns out to be its greatest epistemological problem and, paradoxically, its greatest philosophical contribution, especially due to the incorporation of the notion of common sense in the field of informational semantics. Fodor's computational theory of mind seems to be an undeniable contribution to current models of cognitive science, particularly considering his idea of informational encapsulation. Likewise, the notions of semantic and referential semantic memory, and their possible implications in the domain of Artificial Intelligence, currently constitute an important legacy of the works of the American philosopher. His suggestive observation that there is no such thing as constitutive conceptual truths seems to tip the board toward conceptual intuitionism for now. As long as something more substantive is not evident, the marriage between common sense and belief system, turns out to be the greatest philosophical triumph of 'Citizen Fodor'.

Keywords

Science, cognition, philosophy, information, mind, psychology

Suggested citation: Tillería, Leopoldo (2021). Jerry Fodor's Philosophy of Mind. *Sophia, colección de Filosofía de la Educación*, 30, pp. 151-171.

* Social worker, journalist, Master and Doctor of Philosophy from the University of Chile. Academic and researcher in the Information Technology and Cybersecurity Area of INACAP, Temuco. Research lines: Philosophy of technology, Aesthetics, and Philosophy of religion.

Resumen

El artículo presenta una aproximación a la filosofía de la mente de Jerry Fodor, y se detiene en el problema que representa su tesis del naturalismo computacional. Se enfatiza en las diferencias ontológicas entre los sistemas modulares de entrada y los sistemas centrales de esta máquina mental. Mediante una revisión de las principales críticas al programa fodoriano, se concluye con la idea de que su innatismo conceptual resulta ser su mayor problema epistemológico y, paradójicamente, su mayor aporte filosófico, en especial por la incorporación de la noción de sentido común en el campo de la semántica informacional. La teoría computacional de la mente de Fodor parece ser un innegable aporte a los modelos actuales de las ciencias cognitivas, en particular, considerando su idea de encapsulamiento informativo. Asimismo, las nociones de memoria semántica y semántica referencial, y sus posibles implicancias en el dominio de la Inteligencia Artificial, constituyen en la actualidad una importante herencia de los trabajos del filósofo estadounidense. Su sugerente observación acerca de que no existe algo llamado verdades conceptuales constitutivas, parece dejar por ahora el tablero inclinado hacia el intuicionismo conceptual. Mientras no se evidencie algo más sustantivo, el matrimonio entre sentido común y sistema de creencias, resulta ser el mayor triunfo filosófico del 'ciudadano Fodor'.

Palabras clave

Ciencia, cognición, filosofía, información, mente, psicología.

152



Introduction

Jerry Fodor's work seems impossible to classify from the point of view of his theoretical register. His work focused on the problem of how something called the mind defines a certain type of belief, based on a certain class of mechanisms, has been loved and hated by dozens of cognitive scientists, psychologists of one creed or another, and philosophers of the mind. As Welsh (2016) states, if Fodor's claims are true, then they have profound implications for cognitive science, linguistics, psychology, the philosophy of language, and the philosophy of mind. Fodor's thesis that the mind has a computational architecture, that is, that it would be composed of systems, functions, and properties: it is by far his greatest legacy almost three years after his death. The many times - heightened controversy - between philosophy of mind, psychology, and cognitive sciences, seems to be resolved in the position of the philosopher that there is no way to explain the functioning of the human mind, other than by resorting precisely to a computational approach. In Fodor's own sayings (1986), mental activity should be considered "as a set of formal operations that deal with symbols or representations" (p. 12).

His project, by the way, has consisted of an incisive critique of the teleosemantic program in the philosophy of mind and evolutionary psychology, especially, thinks Fulda (2015), regarding the notion of 'selection of explanations'. For Fodor, it is about determining what is the nature of 'that' which puts the individual in representational contact with

the environment, and, above all, of how it is that which contacts him/her acquires a certain type of meaning. In this way, he examines the great problem of semantics associated with mental representations. In order to functionally specify the concept of mind, Fodor resorts to the literal thesis that mental operations follow a computational mechanism, even if only partially. This means that, once a certain flow of data or stimuli has been computed, it is the central systems of the mind (which, unlike the input and output systems, are not modular) that must execute the reasoning and belief permanence processes. Regarding the latter, his theory seems to leave the field open to the hypothesis about the true nature of these central “endowment” systems of meanings (the semantic “hard drive disk”).

Now, everything indicates that Fodor always presented his representational thesis as an investigation and that at all times, he was aware of the limitations of his theoretical project. A final passage from *Modularity of the Mind* is sufficient as an example: “ We have, to put it bluntly, no computational formalisms that show us how to do this, and we have no idea how such formalisms might be developed.” (Fodor, 1986, p. 177). In this sense, the image that Domingo (2003) outlines about the Fodorian thesis is eloquent in itself:

A nice metaphor to use, we could say that the image of the mind derived from here would be comparable to one of these Swiss army knives in which each of its deployable gear would be designed according to specific purposes (p. 565).

This article is organized as follows: first, what might well be called Fodor’s ‘theory of mind’ is outlined, attempting to emphasize the philosophical status of his project. In the second place, and already fully in its computational theory, the architecture of the input systems is reviewed, to continue in the third section with the specification of the ‘big cabinet’ of mental processes: central systems. The fourth section addresses the problem of the relationship between Fodorian semantics and mental representations. In the end, some of the main objections to Fodor’s program are discussed.

The idea of mind in Fodor’s philosophy

It is necessary to start by breaking down some myths about the Fodor program. The first of them is the one that assumes that because he is a philosopher of the mind, his ideas are only of an ontological rank, and that, therefore, his hypotheses or assertions would play better in the field of metaphysics,

if not in that of fiction literature. Contrary to this objection, Fodor's realistic theory of mind, in particular, that which refers to its modularity, strives to discuss, based on what could be called an open program of psychological research, the a posteriori foundation of intelligent activity in human systems. Undoubtedly, his thesis on the mind collides with a host of cognitive prejudices, especially those of a Quinean or phrenological stamp, which would indicate that the problem of essence or, if you like, of the identity of the mind, is merely philosophical, and therefore should be returned to the Platonic foundations for its resolution, or it is exclusively psychological (or psychiatric or phenomenological), and, in such a case, philosophy would have nothing to do in this kind of 'guild' of experts in cognitive processes.

Of course, this is not the case. There is enough evidence to indicate that the philosophy of mind has emerged precisely from fundamental questions about our states of consciousness, the nature of concepts, belief systems, and mental events, among others. Gilbert Ryle himself, starting from his *The Concept of Mind* (1949), has made a cardinal contribution to the philosophy of contemporary mind, based on his conjecture of the influence on it of the "philosophy of common language" (Botero, 1992, p. 63). Fortunately, and as De Brigard (2017) observes, the increasingly active presence of cognitive sciences, the philosophy of consciousness, and phenomenal consciousness in the current concert of the philosophy of psychology, provides a healthy quota of measure around the relationship between epistemology and the haunting question about the mind.

A first point that should be anticipated about Fodor's thesis is that mental functioning is basically an unconscious activity (a question derived from the possibility of 'access' only to the peripheral processes of the 'general computer': the input systems and output). In other words—and this is probably the question that differentiates Fodor's cognitive ontology from the endeavors of scientific psychology around the problem of mind—Fodorian theory leaves aside the fundamental concern with the problem of consciousness, and it focuses specifically on the spontaneous processes of information analysis and internal representation. What is presented, then, is a philosophy of the psychology of the mental faculties, whose central idea comes to be that of modularity.

Curiously, Fodor's conceptual innateness is, if anything, closer to empirical psychology than to the metaphysics of the mind itself. The observation of Subía and Gordón (2014) seems entirely accurate to focus the Fodorian arguments in their modular perspective:

Empirical psychology, on the one hand, strives to fragment mental structures, while the psychology of the spirit clings to the metaphysical



explanation as an expression of idealism, which resurfaces and does not admit another form of explanation, even less, a that is related to anatomical structures (p. 81).

The orthodox mentalist doctrine assumed by Fodor (1986) can be summarized in the following statement: “It seems obvious that you need mechanisms to put what you know into action; mechanisms that function to bring the organization of behavior into conformity with the propositional structures that are cognized” (p. 29). Content and mechanism are, then, the fundamental determinations of the notion of mind as a functional device. Fodor’s hypothesis also adopts a birth certificate adhering to physicalism and intentional realism. His approach is physicalist, insofar as he declares that mental processes suppose the preexistence of physical processes capable of causing said phenomena. Thus, the American philosopher believes that the laws introduced in any special science are determined by the existence of more basic physical processes that implement such laws (Rengifo, 2008). Its intentional realism, in turn, supposes that position that maintains that the natural order shows certain intrinsic intentionality, in such a way that, connected this intentional realism with the problem of the intentionality of the content, it would have, according to Vallejos (1990), that the first would make possible the formulation of the conditions under which the content of a representation can be determined in naturalistic terms.

Put schematically: the architecture of the mind in Fodor’s program defines three types of components. On the one hand, there are the transducers (sensory and motor), which interface the mind with the outside world through a purely physical interaction, which supplies, or is influenced by, computationally suitable information codes, directly linked to proximal properties of the stimulus of the environment (García-Albea, 2003). Secondly, there are modular input and output systems, with the ability to generate inferences and trigger representation states in an automated manner. Finally, with García-Albea (2003), there are the central systems (the heart of the computer), that is, computer systems of general purpose, interactive and sensitive to global properties of all the available information. For Fodor (1986), this functional taxonomy ultimately determines three classes of exclusive psychological processes. However, he is also extremely cautious about it. Writes Fodor (1986):

Since the trichotomy is not exhaustive, it is left wide open that there may be modular systems that do not subserve any of these functions. Among the obvious candidates would be systems involved in the motor integration of such behaviors as speech and locomotion. (p. 69)

In short, and as Gomila (1991) thinks, Fodor defends the existence of a representational system in the mind, whose properties make it similar to a language: productivity, systematicity, intentionality, referentiality: “In the sense that this representational system is innate, and very rich — in fact, maximally rich, since it includes all uncompounded concepts” (p. 36). In the following section, the ontology of modular systems will be described, that is, the inferential mechanism that generates the necessary syntax to encode the information supplied by the transducer systems in mental representations, endowed with a minimum semantics for the operations of the central processor.

Modular systems: autonomy and inferential representationism

156



Based on the above, Fodor (1986) will argue that input systems imply, by definition, a certain syntax, so that what enters such systems seems to be assured in relation to the possibility of a certain representation of the world, until the point that it is finally accessible to thought. Such systems specialize, Murphy (2019) comments, in performing quick calculations, because they are not influenced by information from other modules or the central system. The complexity of the spectrum of accessible information (for example, the range of properties of such informational objects) determines, following Fodor (1986), the assumption of the existence of highly specialized computational mechanisms in the task of generating hypotheses about —precisely— the distal sources of proximal stimulations.

These hypotheses are housed in an explanatory conglomerate that Fodor calls Representational Theory of Mind (RTM). “One generates conditions for the semantic evaluation of an attitude when setting a context for the samples of certain symbols: symbols that, together, constitute a system of mental representations” (p. 145). In other words: the content of mental representations (what is really understood when it is understood) is determined not by the intentionality of certain representational attitudes, but, says Fodor (1994), by the interpretation of a certain primitive nonlogical vocabulary, which is a condition for representational syntax. As the philosopher Liza Skidelsky (2006) observes, such modular systems alternate autonomy and inferential representations:

According to what I will call a substantial conception of modules, these are autonomous psychological mechanisms that are designed for the processing of cognitive information. As computational-inferential me-

chanisms, they have as premises-inputs the transduced representations of nearby stimulus configurations and as conclusions-outputs the representations of the nature and distribution of remote objects (p. 85).

Another way of understanding this kind of contextual determination on the representations that are computed by the input systems is the idea, thinks Fodor (1986), that these modular systems operate on eccentric domains because the computations they execute turn out to be necessarily idiosyncratic. Indeed, the condition of specificity of some of these peripheral systems is evident, precisely given the multivariate nature, so to speak, of the incoming stimuli. In this way, the modules are pre-established, that is, not formed, or, what is the same, essentially primitive and, rather than fast, automatic. When such modules are activated by the corresponding stimulus, they respond by producing data about it, which, being automatic, are neither elaborated nor informed by central cognitive processes (Bacáicoa, 2002). It is this condition of automaticity that, as someone says, “forces” each module to always be an available receptor for all types of stimuli in strict sensory correlation. Fodor (1986) emphasizes:

In short, the operation of the input systems appears to be, in this respect, inflexibly insensitive to the character of one’s utilities. You can’t hear speech as noise even if you would prefer to. (p. 84).

Seen this way, what is connected or disconnected is nothing but the access of certain input systems to certain central systems of the mental machine, specifically to its thought structure. In this way, the possibility of disconnection/connection —its operational flexibility— will be characterized by the rationality with which the organism exploits the resources provided by its representational system. In Fodor’s explanation (1985): for the intelligent use the individual makes of his/her internal representations. This strictly functionalist conclusion, that the final processing of stimuli that are transformed into representations depends on a certain disposition of rationality in each organism, radically places Fodor’s research in line with the harshest physicalisms.

In any case, it should be clarified that the input systems, when they produce the “initial” hypotheses, do not generate real beliefs. Instead, according to Meyering (1994), they only pose perceptual hypotheses about the external world. That is, on the basis of these input systems the central system finally ‘decides’ what to believe. Consequently, Meyering (1994) adds, what Fodor calls ‘perceptual belief fixation’ is the exclusive prerogative of the central system. It is clear, as García-Albea (2003) lucidly warns, that Fodor’s original motivation was to explore the possibility of an alter-

native proposal to the dominant *New Look* approach, establishing a precise demarcation between the most basic aspects of perceptual processing and Higher-order cognitive processes, on which depends, among other things, the fixation of one's perceptual beliefs and, in general, all kinds of non-demonstrative inferences.

Central isotropic systems

Care must be taken, then, in specifying the relationship established by Fodor between modular entry systems and central systems. Of course, there is a fundamental informational connection regarding the semantically catalytic role of the latter. Such connection refers to the fact that the information is part of the same frame of reference of the individual; so to speak, it is about the same line of thought, which, moreover and generally, is computed in the different mental modules in the same code or language system. However, there are substantive differences from the structural-functional point of view itself. While the input systems appear to follow a certain top-down verticality, the central processing systems are distinguished by a more holistic horizontal architecture. Thus, García (2005) states:

Central systems, by contrast, have a highly distributed neuronal base in the brain that is very difficult, impossible, Fodor claimed, to identify; they are more open to learning; they operate slowly; as they are not encapsulated, they receive information from various modules; they integrate information from the outside world with the information recorded in the individual's memories. The core systems are thus holistic and not encapsulated (p. 7).

The central systems have the essential function of providing certain mental representations in relation to the data that the input systems provide. Such systems elaborate, suggests Fodor (1986), probable hypotheses about the state of affairs in reality. These hypotheses, which in the long run constitute belief systems about certain contextual circumstances, seem, in turn, to confirm that Fodor's research suggests a new theory of rationality (innate and physicalist, as is hinted at), which attempts to account for the representational mental processes. In the words of Hermida (1993):

The explicit representation is required of the occurrences of thoughts involved in a mental process, that is, of the data structures in Game ('data structures'), without it being required of the transformation pro-



grams and rules (corresponding to the rules of thought) that could well operate without being explicit (p. 364).

These representational processes constitute the crux of Fodor's TMR, whose 'genome' is the language of thought. The product—it could be said in computer terms, the 'machine language'—corresponds to the propositions expressed precisely by means of these representations. Thereof, semantics and syntax "assemble" their rationality game in a computational way. In short: syntax plays the role of mediator between the causal/formal properties of representations and the semantic properties of the propositions expressed by such symbols, thus allowing the syntactic processes that operate on the representations to preserve the truth of said propositions (Rengifo, 2008). This kind of 'hard disk', which the central systems appear to be, is defined by their state of hyperconnectivity, or if you prefer, by their exponential combinatorial capacity. Fodor explains (1994):

This is explained [the correlation between types of belief states and distinctive intentional object and causal role] immediately if it is assumed that belief states are somehow constituted from elements and that intentional objects and causal roles of each of the states depend on what elements it contains and how those elements are put together (p. 208).

It is the same as saying that central systems allow thought and behavior to be connected. Individuals—it is evident—generate their behaviors because they have a representational system. The fact of deciding such behaviors, observes Medina (2010), is a computational process. As Fodor himself (1986) argues: "(...) there must be relatively nondominant (i.e., domain-inspecific) psychological systems which operate, inter alia, to exploit the information that input systems provide. Following the tradition, I shall call these central systems" (p. 146). In this way, and as a possibility of semantically associating thought and behavior, computational processes must necessarily be symbolic and formal. Symbolic, because they are defined with respect to certain representations, and formal, as Rengifo (2008) reasons, because they are applied to representations based on their higher-level physical properties.

The operational formulation of these nonspecific central systems (this is a decisive difference with respect to the input systems, which are rather defined as encapsulated with respect to a certain data referent with which they contact), would be more or less this: from the inputs from the work of the input systems, the central systems proceed to the fixation of beliefs by means of non-demonstrative inferences. This process of fixing beliefs occurs from two supply lines: one, the provision of data through



the input systems, and the other, the information available and stored in memory. By treating these two types of information, emphasizes Fodor (1986), central systems generate the most probable hypotheses about the state of affairs in the world.

Now, Fodor's intentional realism starts from certain principles that the philosopher considers unobjectionable, for example, the claim that mental states do translate semantics of some kind. Fodor argues (1994):

Roughly stated, we are arguing about whether they have *combinatorial* semantics, the kind of semantics in which there are (relatively) complex expressions whose content is determined, in a regular way, by the content of their (relatively) simple parts (p. 194).

Precisely, this informational semantics responds to the computational structure of a mind that develops certain cognitive processes in relation to the state of things that surround it. Such processes are computational to the extent that their representational characteristic is based on symbols that allow their manipulation. From this, it follows that the main characteristic of the mind is that it is a symbolic representational system within which a certain type of thought takes place, which is expressed through also symbolic computations (Medina, 2010).

Fodor's RTM, consequently, literally becomes a theory of cognitive systems, if the functions of the aforementioned input systems and central systems —periphery and heart, respectively— of this thinking machine are understood as such. Following this reasoning, Fodor (1986) establishes a kind of final conjecture that seems quite obvious: “if the analysis of the inputs is modular and the thought is quinean/isotropic, our brain will consist of a stable neural architecture associated with perception and language, but not thought” (p. 164). This seems to be easily explained by adducing a kind of line difference between thought and behavior (hence the instability of thought referred to), or, if you like, between mental representation and propositional attitude. It is what Martínez (1995) calls the psychology of sense or mental innateness:

On the other hand, it is characteristic of the psychology of common sense or popular psychology to be a psychology of belief/desire, as Fodor emphasizes in *Psychosemantics*, that is, a set of psychological knowledge that human beings possess de facto about mental processes in themselves and in their relation to behavior, in which the concepts of desire and belief play a central role, in such a way that we explain and predict people's behavior assuming that they act according to their desires and beliefs (p. 369).



From this, stems what Fodor (1994) calls the language of thought (LOTH) is none other than the internal code that the RTM requires to make its levels of propositional semantics explicit. The LOTH establishes that to have a thought is to be related to a formation of representations. Presumably, Fodor (1994) points out, having the thought that John loves Mary is, ipso facto, having access to the same representations and representational structures required to have the thought that Mary loves John. However, as will be seen in the section that follows, Fodor's functional-physicalist explanation considers that central processing systems must be understood from an architecture different from the mere purely computational arrangement with which modular systems had been defined until now. peripherals. These central systems lack specific contents that depend on specific neural structures.

For Fodor (1986), all the available data indicate that the central problem-solving process rests on equipotential neural mechanisms, a conclusion reached by assuming that the central cognitive processes themselves are fundamentally quinean and isotropic. Hence, according to Rey (2018), possible thoughts have, for Fodor, a recursive logical structure, which allows people to be able to think a potential infinity of increasingly complex thoughts, from which it would even be doubtful that there is any finite upper bound type.

Fodor's semantic model

Fodor's non-reductionist physicalism (token physicalism) is still explanatory in terms of the causations that it seeks to determine for the cognitive processes and output behaviors of each individual, known until now as propositional attitudes. Precisely this approach suggests the idea of a philosophy of mind that recognizes conditions of causation of different kinds for mental states. Fodor explains (1994):

It is characteristic of the common-sense psychology of beliefs and desires—and therefore of any explicit theory that I am willing to regard as a theory that vindicates the common-sense psychology of beliefs and desires—that it attributes content and causal powers to the same mental things that are considered to be *semantically evaluable* (p. 32).

It is, in short, a semantic recursion mechanism, which is referred to as a reasoning model by means of which belief systems become what they are (truthful formulations about the world) according to the meaning assigned to them. Such recursiveness, undoubtedly, is determined by

a reassessment of the aforementioned propositional attitudes. This claim seems to be articulated around the collection of experiences and intuitive testimonies on which it will be reflected:

To highlight the virtues of this psychology of common sense: its operability and credibility (how often it goes right), its depth and its essentialness (how much we do depend upon it) (Hermida, 1993, p. 361).

Contrary to what could be assumed with respect to a functional-computational project, Fodor's thesis about mental representations does not rise to the top of a bio-informatic epistemology, but rather includes the assumption of the possibility of a connection between common sense and beliefs. A passage from Fodor (1994) himself should be quoted: "We have no reason to doubt—in fact, we have solid reasons to believe—that it is possible to have a scientific psychology that vindicates common sense explanations through beliefs and desires" (p. 37). Indeed, this sort of common-sense rationality is expressed in a mechanism in which a syntactically regulated mental machine can determine, with a high probability of adjusting to the environmental conditions, the semantic properties of the symbols it uses.

Recursion and rationality are, then, conditions of possibility so that the mind can define a certain suite of expected behaviors according to the type of administered semantics. In this regard, Cárdenas-Marín (2016) maintains, in a properly Fodorian line:

According to Kripke, names do not have meaning, so the reference is necessary for it to be pointed out, understood. It should be noted that these referential and significant elements seem to be mostly conventional elements around the name (p. 117).

In this way, the computed sequence is composed of: a) mental representations, b) mental processes and c) intentional causations. On this point, Fodor (1994) mentions: "From this, it follows that the samples of attitudes must correspond with samples of mental representations when they—the samples of attitudes—are episodes in mental processes" (p. 48). The motto, he says, should be: "There is no Intentional Causation without Explicit Representation." This means that having a certain repertoire of behaviors with respect to the environment is being in a certain relationship with an internal mental representation.

Now, this attitude-representation relationship seems to corroborate a perfect computational aspect. Mental states, Fodor (1985) thinks, are relationships between organisms and internal representations, and such



states, causally interrelated, succeed each other according to computational principles that are formally applied to such representations. Hence, the data processing of this operations center is the domain of the mental representations that transform the causal processes (the information from the input systems) into certain propositional attitudes as response states with respect to the environment. The fact that the RTM shows how intentional states have causal properties is, Fodor (1994) asserts, the most disturbing aspect of common-sense intentional realism from a metaphysical point of view. This concern is reduced, finally, to the idea that the concepts—that is, the informational prototypes that make certain codes entered in this resonance box mean certain things and not others— are mental artifacts physically instantiated. As Rellihan (2009) reflects:

When Fodor asks for a “physical” or “mechanical” account of mental processes, he is asking for an account of their implementation mechanisms. Psychology is a special science and, in Fodor’s analysis, for each special science law of the form “Fs causes Gs,” there is something to tell about the “lower-level” law or laws—the mechanisms of implementation—in by virtue of which Fs causes Gs. Thus, just as the laws of thermodynamics are implemented by the mechanisms of statistical mechanics and the laws of inheritance by the mechanisms of molecular biology, then the laws of psychology must be implemented by their implementation mechanisms, whatever those mechanisms may be (p. 314).

163


Recognizing that concepts can be acquired, even if they could not be learned, Fodor (2008) explicitly admits that the connection between experience and the concept it engenders can be completely fortuitous: a concept could emerge “through surgical implantation or swallowing a pill, or hitting the head against a hard surface” (cit. in Antony, 2020, p. 46). In further specification, Fodor (1998) adds: “I affirm that acquiring a concept is to link up nomologically with the property that the concept expresses” (p. 177). The conditions of confirmation of a concept, insists Fodor (1998), are not among its essential properties: “The confirmation [of a concept] is an epistemic relationship, not a semantic relationship and, generally, it is mediated by theory, therefore, is holistic” (p. 47).

It has been said that the principles that apply to mental representations, based on the sequence of mental states, are formal in nature. Fodor, explains Ramírez (2019), intends to offer a robust theory that can face the problems raised around the erroneous representation and the problem of disjunction, establishing asymmetric causal dependency relationships between the properties, which are enough to cause a concrete mental representation. It remains, then, to elucidate that of the formal applied to

this computational theory of the mind. For Fodor, computational processes are formal, in the sense that they apply to representations by virtue of their higher-level physical properties. That is, formal operations are specified without referring to the semantic properties of the representations, such as truth, reference, or meaning (Rengifo, 2008). In other words, the fact that Fodor's computational theory is formalistic, means, asserts Zumalabe (2014), that his information processes are involved in a descriptive relationship with the notion of algorithm, defined in terms of processes that operate on representation.

This means that Fodor's RTM, bypassing such swampy waters for cognitive sciences as semantic holism, embodied realism, and connectionist models, postulates, through its formalization in a computational theory of mind, the specification, Hermida (1993) argues, from "a computer between the sensory systems and the belief box to, through a certain system of computational calculations, transform the occurrence of a psychophysical property into an instance of an internal mental symbol" (p. 369). In any case, as Fodor himself mentions, the multiple realizability of the mental (that is, the physicalist causation of mental states) is an empirical hypothesis. In this sense, Bermúdez and Cahen (2020) point out, its plausibility will depend on the evidence that can be had for the multiple/real realization of the mental.

164



Criticism of irrationalism? by Fodor

The Fodorian theory of mind has faced innumerable theoretical, epistemological, and methodological objections. Among the most forceful criticisms, there is the one known as that of the 'constructivist fallacy', which, as can be expected, shoots into the center of Fodor's conceptual innateness. It is convenient to start with this "first line" of criticism.

The 'constructivist fallacy' argument is directed at Fodor's idea that states that the concepts (at least the fundamental ones, or to put it colloquially, the 'trunks') with which the mind works are innate. Such reasoning, indicates Gomila (1991), known as the 'Fodor paradox', implies affirming that to learn a concept it is necessary to have it previously instantiated, which, to avoid the threat of infinite regression, translates into all non-compounded concepts are innate. Fodor's argument is presented with such radical abstractionism that it would certainly make it difficult, to say the least, to establish a nomological or even a properly physicalist foundation in relation to the emergence of such concepts. In a kind of

anticipation of an expected counterfactual attack, Fodor (1998) asserts: “Informational semantics is a *content* theory and these needs could be seen as *metaphysical* rather than semantic” (p. 111).

Now, and as Gomila (1991) points out, with the argument of conceptual innateness Fodor seems to deny, above all, common sense aspects, such as the diversity of conceptual repertoires depending on the context and the creativity of new concepts, which could hardly be given up. This impasse seems to be the same as that presented in relation to the problem of the epistemological representation of an individual object. Thaliath observes (2019):

The existence of the concept implies an ontological aporia against material objects, as well as mathematical objects, to which Plato seemed to attribute an intermediate ontological state — that is, an ontological state between the meanings and eternal ideas. This seems to have resulted from the ambiguity of an epistemic reference (p. 146).

A second element of Fodor’s theory that has been the target of cognitivist criticism is one that points to the heart of the mental reasoning mechanism: the problem of syntax. If his computational mental theory has been described as a system that depends, both on the causal/nomological relationships established by mental representations with the things that fall under them (Rodríguez, 2006), and on the process of data coding by a ‘central brain’, then, the process of generating mental states and representations depends, critically, on what could be called, in a way, an ontology of syntax.

However, Fodor’s project does not develop a detailed ontology of the syntactic mechanisms that formally determine representational operations. In other words, Fodorian theory fails to specify the syntactic formalism required to account for an informational-computational semantics. If the syntax has been changing with evolutionary processes, or if it is rather a universal syntax, it is clear, in both cases, that such syntax ceases to have a logical parallel with Fodorian computational theory. In fact, in the first case, it would force to establish the mechanisms capable of sustaining these changing computations, which would make the model, if we follow Cela-Conde and Marty (1991), literally into a kind of syntactic coven.

If Fodor’s departure from this objection about the syntax of his representational system is the same thesis of the syntactic-semantic condition that constitutes the representational mechanism —his commented conceptual innateness— it would seem that the argumentation is on the

verge of a tautology. Certainly, if the combinatorial/compositional syntax consists, following Skidelsky (2012), in that complex mental representations are constructed from atomic mental representations by means of syntactic rules, then the 'syntactic chain' is converted, *a fortiori*, in a mechanism as innate as the very fundamental concepts described by Fodor.

A third chord from which the criticisms of Fodor's theory come is, to put it in the nomenclature of the philosophy of mind, the scope of its external justification. Domingo (2003) affirms that Fodor's thesis, at least the one framed in his latest works, 'does not pass the test' of a fourfold justification: phylogenetics, ontogenetics, neuropathology, and neurocerebral. This would mean that Fodor's modular theory fails to demonstrate the expected compatibility with these fields of scientific psychology. From the phylogenetic and ontogenetic planes, Fodor's work is shown, from the beginning, Domingo (2003) notes, reluctant to consider any determination regarding a supposed evolutionary linearity between the brain and cognitive structures. Regarding the neuropathological and neurocerebral fields, Domingo's comment (2003) is no less benevolent:

Fodor does not record the information emanating from research on brain injuries and the motor, sensory, cognitive, linguistic, and emotional consequences they produce, nor does he contemplate the holistic principle of the plasticity of the human brain that allows, within certain limits, the recovery of the functions lost due to injury in other undamaged areas of the brain (...), as regards the neuroanatomical and neurobiological level, the consequences (...) are much more devastating. According to the studies by Damasio (1996), Gazzaniga (1996) or Le-Doux (1999), the brain not only has an extremely intricate network of neural connections that flow into the frontal lobe (...) from practically all regions of the brain, but they have also discovered in it the existence of a double loop-shaped circuit through which the "sensory systems", those that are the best candidates for the supposed mental modularity, would have access almost directly to the basal ganglia and especially to the amygdala (both privileged organs to configure the somatic markers that collect the background state of the body), for all of which two of the basic principles prescribed by Jerry Fodor seem definitely invalidated: informational encapsulation and modular independence (p. 570).

In any case, Fodor does not pass up this criticism. His idea of neural architecture is congruent with the computational mental structure that he defends. Thus, for example, Chow (2016) comments, neural structures and stable patterns of connectivity and information flow would be seen in the parts of the brain dedicated to peripheral processing (perception,



language, motor control, etc.), but this would not be seen in the brain centers involved in general reasoning. Instead, instantaneous and unstable neural connectivity would be seen in the association cortices that ‘seems to go in all directions’.

The ‘last cartridge’ of criticism that will be addressed in this brief review, refers to the problem of the naturalistic explanation of the functioning of mental processes. At this level, there seems to be a greater brand ambiguity, if, on the one hand, mental processes are made dependent on the syntax of mental representations, and if, on the other, these same processes are made dependent, in relation to the fixation of certain beliefs, of the context. Granting that computationally only the possibility of the first premise can be verified, Igoa (2003) will affirm that the resulting paradox is expressed in the need to restrict computational explanations only to the functioning of the modules. Igoa argues (2003):

In the case of non-modular systems, on the other hand, the syntax of the representations they handle (and, therefore, the computational processes) is not enough to explain the work they perform. If the entire human mind were modular, then repeating the title of Pinker’s book, we might know *how the mind works*, but if only a part of the human mind is, we will only know, as Fodor claims, that the mind probably doesn’t work. that way (that is, computationally) (p. 534).

That is to say: Fodor’s modular theory seems, at least from the point of view of its computational ontology, to fail the test of a minimal attempt at syntactic-semantic falsification. However, and as an indispensable counterargument to what could be seen as a definitive epistemological critique, Fodor’s program has never been intended to convert cognitive psychology into the psychology of knowledge. Moreover, Quiroga (2010) supposes, one of the fundamental assumptions of Fodor’s work has always been to postulate a computational theory of mind, leading to a type of methodological solipsism as part of research in cognitive psychology. The question is whether its conceptual innateness, which could also translate into the inability of the philosophy of mind to identify the conditions for the generation of the deepest mental mechanisms)— its ‘extreme epistemological caution’ (Domingo, 2003))— leaves or not Fodor in the field of irrationalism.

Conclusions

From what has been outlined, where Fodor’s approach is based on an almost infinite series of ‘isms’, such as innateness, computationalism, for-

malism, conceptualism, and a long etcetera, we will try to exempt him from a certain part of the criticisms which he bravely faced. Of course, there are a relevant number of objections (epistemological, methodological, ontological) to which Fodor's theory has failed to answer satisfactorily. That of his computational naturalism seems to be the most severe. However, as Cela-Conde and Marty (1991) point out, the theses of 'Citizen Fodor' (as both authors call it in relation to the criticism they make of his computational model, which, from the point of view of possible intelligences, 'theoretically' would have aborted his existence as a simple 'citizen'), have been indispensable in generating a healthy philosophical discussion about mental states, intelligence and the role of computational models in cognitive processes. Bennett and Hacker (2003) themselves suggest, in Braun's (2007) statements, that the attribution of psychological predicates to the brain is primarily a philosophical question, and not so much a neurological one, since it is essentially a conceptual question.

Jerry Fodor has clearly missed the philosopher's stone of the mind. His modular-computational theory, however, seems to be an unquestionable contribution to current models in the philosophy of mind, in particular regarding his idea that informational encapsulation is the primary characteristic of mental modules (Bacáicoa, 2002). This thesis, as it were, of the compartmentalization of information by modular encapsulation, has at the same time become the cardinal breakpoint with the world of cognitive sciences, and that, without considering the dispute over the critical role of the isotropy in higher-level cognitive processes. Paradoxically, Fodor's project leaves several pending tasks to the cognitive sciences themselves. First, the challenge of finding empirically more reasonable explanations for the workings of the mind, which do not fall, is obvious, in evolutionary or biologic explanations. Secondly, the task of solving the problem of the origin of concepts, a question that should not imply, as expressed by Fodor (1998) himself, the blind acceptance of informational atomism.

On a somewhat different side, the notions of semantic memory and referential semantics, as well as the potential implications of both structures in the field of Artificial Intelligence, seem to owe much of their preponderance to Jerry Fodor, now in the domain of philosophy of technology. The philosopher's brief observation that there is no such thing as constitutive conceptual truths, and that consequently there would not be any kind of definitions either, leaves the board tilted towards conceptual intuitionism for now. Apparently, and while something more substantive is not evidenced, the marriage between common sense and belief system turns out to be the greatest philosophical triumph of 'Citizen Fodor'.



Bibliography

- ANTONY, Louise
2020 Not rational, but not brutally causal either: A response to Fodor on concept acquisition. *Theoria. An International Journal for Theory, History and Foundations of Science*, 35(1), 45-57. <https://doi.org/10.1387/theoria.21031>
- BACÁICOA, Fernando
2002 La mente modular. *Revista de Psicodidáctica*, (13), 5-30. Recuperado de: <https://tinyurl.com/yyoa5wb5>
- BERMÚDEZ, José Luis & CAHEN, Arnon
2020 Fodor on multiple realizability and nonreductive physicalism: Why the argument does not work. *Theoria*, 35(1), 59-74. <https://doi.org/10.1387/theoria.20772>
- BRAUN, Ricardo
2007 El futuro de la filosofía de la mente. *Persona* (10), 109-123. Recuperado de: <https://tinyurl.com/y3tdc8l2>
- CÁRDENAS-MARÍN, William
2016 Reflexiones sobre el nombrar a partir de la comprensión del lenguaje de Saul Kripke. *Sophia, colección de Filosofía de la Educación*, 20(1), 103-116. <https://doi.org/10.17163/soph.n20.2016.04>
- CELA-CONDE, Camilo & MARTY, Gisèle
1991 El ciudadano Fodor no existe. Estructura y función en los procesos cognitivos. *Psicothema*, 3(1), 231-241. Recuperado de: <https://tinyurl.com/yyhltq5l>
- CHOW, Sheldon
2016 Fodor on global cognition and scientific inference. *Philosophical Psychology*, 29(2), 157-178. <http://dx.doi.org/10.1080/09515089.2015.1013208>
- DE BRIGARD, Felipe
2017 El problema de la conciencia para la filosofía de la mente y de la psiquiatría. *Ideas y Valores* (66), 15-45. Recuperado de: <https://tinyurl.com/y7kqxtm8>
- DOMINGO, Josep María
2003 El proyecto modular de Jerry Fodor (o sobre el porvenir de otra ilusión). *Anuario de Psicología*, 34(4), 505-571. Recuperado de: <https://tinyurl.com/y3kk7ret>
- FODOR, Jerry
1985 *El lenguaje del pensamiento*. Madrid: Alianza.
1986 *La modularidad de la mente*. Madrid: Morata.
1994 *Psicosemántica. El problema del significado en la filosofía de la mente*. Madrid: Tecnos.
1998 *Conceptos*. Barcelona: Gedisa.
- FULDA, Fermín
2015 A mechanistic framework for Darwinism or why Fodor's objection fails. *Synthese*, (192), 163-183. <https://doi.org/10.1007/s11229-014-0557-4>
- GARCÍA, Emilio
2005 Teoría de la mente y desarrollo de las inteligencias. *Educación, desarrollo y diversidad*, 8(1), 5-54. Recuperado de: <https://tinyurl.com/y6f98dra>
- GARCÍA-ALBEA, José
2003 Fodor y la modularidad de la mente (veinte años después). *Anuario de Psicología*, 34(4), 505-571. Recuperado de: <https://tinyurl.com/y6q57owk>



- GOMILA, Antoni
 1991 El innatismo de Fodor: Consideraciones críticas. *Estudios de Psicología*, (45), 35-48. Recuperado de: <https://tinyurl.com/y67ret8t>
- HERMIDA, Pablo
 1993 La psicosemántica de Jerry A. Fodor. *Endoxa Series Filosóficas*, (1), 359-372. Recuperado de: <https://tinyurl.com/ycm8ft9h>
- IGOA, José Manuel
 2003 Las paradojas de la modularidad. *Anuario de Psicología*, 34(4), 505-571. Recuperado de: <https://tinyurl.com/y5d8kzyn>
- MEDINA, Nicolás
 2010 Realidad mental en la ciencia cognitiva. *IIPSI*, 13(1), 209-220. Recuperado de: <https://tinyurl.com/y3h6rtlh>
- MARTÍNEZ, Pascual
 1995 Wittgenstein y Fodor sobre el lenguaje. *Anuario Filosófico*, 28(2), 357-376. Recuperado de: <https://tinyurl.com/yynb74t>
- MEYERING, Theo C.
 1994 Fodor's modularity: A new name for an old dilemma. *Philosophical Psychology*, 7(1), 39-62. <https://doi.org/10.1080/09515089408573109>
- MURPHY, Gregory L.
 2019 On Fodor's First Law of the Nonexistence of Cognitive Science. *Cognitive science*, 43(5). <https://doi.org/10.1111/cogs.12735>
- QUIROGA, Lautaro
 2010 *Modelos mentales y razonamiento* [Tesis de grado]. Universidad de Chile.
- RAMÍREZ, Camilo
 2019 Sobre conejos y otras figuras geométricas. El problema de la indeterminación referencial en la teoría semántica de Fodor. *Discusiones Filosóficas*, 20(34), 165-184. <https://doi.org/10.17151/difil.2019.20.34.9>
- RELLIHAN, Matthew J.
 2009 Fodor's riddle of abduction. *Philos Stud*, (144), 313-338. <https://doi.org/10.1007/s11098-008-9212-6>
- RENGIFO, Mónica
 2008 *El rol de la representación mental en la ciencia cognitiva* [Tesis de Magister]. Universidad de Chile.
- REY, Georges
 2018 Remembering Jerry Fodor and his work. *Mind & Language*, 33(4), 321-341. Recuperado de: <https://tinyurl.com/y29tjb4e>
- RODRÍGUEZ, Manuel
 2006 *Sobre conceptos. Atomismo Informacional* [Tesis de Magister]. Universidad de Chile.
- SKIDELSKY, Liza
 2006 Modularidad e innatismo: una crítica a la noción sustancial de módulo. *Revista de Filosofía*, 31(2), 83-107. Recuperado de: <https://tinyurl.com/yycvhp5>
- 2012 ¿Lenguaje del pensamiento o pensamiento en lenguaje natural? En C. Silva y L. Salvatico (Eds.), *Filosofía e História da Ciência no Cone Sul*. Porto Alegre: Entrementes.



SUBÍA, Andrés & GORDÓN, Jacqueline

- 2014 Esbozo crítico sobre las estructuras cognitivas: Génesis del pensamiento científico. *Sophia, colección de Filosofía de la Educación*, 16(1), 71-82. <https://doi.org/10.17163/soph.n16.2014.03>

THALIATH, Babu

- 2019 Lenguaje y referencia. *Sophia, colección de Filosofía de la Educación*, 27(2), 141-166. <https://doi.org/10.17163/soph.n27.2019.04>

VALLEJOS, Guido

- 1990 Realismo intencional y naturalización del contenido. *Lenguas Modernas*, (17), 5-34. Recuperado de <https://tinyurl.com/y4bspbva>

WELSH, Sean

- 2016 Jerry A. Fodor and Xenon W. Pylyshyn: Minds Without Meanings: An Essay in the Content of Concepts. *Minds & Machines*, (26), 467-471. 10.1007/s11023-016-9399-9

ZUMALABE, José María

- 2014 Acerca del carácter representacional de la mente. La mente representacional. *Psychology, Society, & Education*, 6(2), 125-144. Recuperado de: <https://tinyurl.com/y2w8t3mo>

Document reception date: July 15, 2020

Document review date: September 15, 2020

Document approval date: October 15, 2020

Document publication date: January 15, 2021