

# Conceptions of the mind... that do not loose sight of logic\*

Juan José ACERO

ABSTRACT: Which is the relation between logic and philosophy of mind? This work tries to answer that question by shortly examining, first, the place that is assigned to logic in three current views of the mind: Computationalism, Interpretativism and Naï ve Naturalism. Secondly, the classical debate between psychologism and antipsychologism is reviewed –the question about whether logic is or not a part of psychology- and it is indicated in which place of such debate the three mentioned conceptions of mind are located.

Keywords: Logic, philosophy of mind, computationalism, interpretativism, naive naturalism, psychology.

What is there in recent philosophy of mind that a logician could be interested in? In my opinion, one should give to this question a two-sided answer. First, I contend that the place of logic is assigned to significantly depends on what the view of the is like and, in accordance with this point, I will briefly dwell upon three views of the mind: Computationalism, Interpretativism and Naive Naturalism. Secondly, I will consider the way in which logic is integrated in each of these views and show the new guise that the classical Psychologism thesis—whether logic is or not a chapter of psychology—comes in.

## *Minds as machines*

According to Computationalism, mental activity is nothing but the formation and manipulation of symbols, i.e., representations. To perceive, to make a decision, to imagine are computation varieties, results of very simple operations that handle representations belonging to a language of thought (Mentalese in loose talk). A proposal concerning which operations are those was provided in Alan Turing (1937). The essential feature of those operations lays in their only being sensitive to formal properties of Mentalese's symbols. Two mental representations differ as far as they do not share each and everyone of their formal properties. These differences codify differences in their causal roles, i.e., differences in their causal antecedents and/or consequences, in their computational antecedents and/or consequences. Any system able to perform the extremely simple operations he identified, Turing had it, would have the resources to think. Minds are machines. Even more, since mental manipulation is a formal business, minds are syntactic machines.

In delineating the foundations of the Computational view of the mind, logical theory has much to say. The computationalist uses to think that logic has the key to understanding human intelligence and rationality. Mind not only comprises many diverse kinds of states, but a complex system of causal transitions among those states as well: perceptions that give rise to beliefs, beliefs that generate, and are generated by desires, desires that bring about actions, and actions that give rise to new beliefs, desires and

---

\* I am indebted to Emily Jane Moreno Spooner and Manuel de Pinedo for having made the paper's English readable. Manuel and Neftalí Villanueva have also contributed many coments on matters of both form and content.

perceptions; and the whole cycle starts again. However, all this activity is intentional. This means that mental states have intentionality, content; semantics in addition to syntax. It is content that gives mental states part of their identity and explains why causal connectivity is the way we find it. How is it possible that intentionality runs along the web of causal transitions, how is it that causes become reasons—that is the question. Some believe that mental state content is a function of its causal relationships to other states, to sensorial input and behaviour output (Loar 1981). Others extend those links to include real and possible objects and situations belonging to the environment that the mind might be nested in (Harman 1999). A third party (Fodor 1987; 1994) chooses to hold that atomic Mentalese's symbols win content in virtue of their tokens' holding causal relations to properties instantiated by external-world individuals. Mental content, therefore, is a result of adequately combining those atomic meanings by compositionality principles that belong to mind architecture.

Logical theory has had a lot to contribute to make this realm surveyable. Logic's general subject matter is the system of properties of terms and sentences that endows them with the form, i.e., the logical form, that explains why logical truth and contradiction, consistency, completeness, logical consequence and independence, etc. are properties sentences which any language might have or relations they might maintain with each other. Now, Computationalism holds that Mentalese is a language—the language of thought. What we commonly call concepts are but its terms; and what we usually call thoughts are but its sentences. Therefore, in so far as their formal, syntactic, properties and relations serve as semantic properties and relations we are in a comfortable position to understand how it is that intentional properties and causal transitions are correlated to each other. In other words, to see why causal transitions among mental states are computational transitions (Fodor 1980; 1987). Turing hit on the right answer to this question in suggesting, first, that a computational system is a syntactical machine and, second, that it is filled up with semantics, i.e. intentionality, if there is a one-to-one mapping between content and form.

A lot of logical research can contribute to exhibit intentional properties more and more subtly embodied in formal niceties. A well known example, that has meant hard work for logicians in the recent past, is that of non-monotonic logical consequence: a sentence  $p$  follows from a set of premises  $A$ , while  $p$  does not follow from  $A \cup B$ . From a computationalist point of view, non-monotonic reasoning is characterized by the fact that those causal links that make you believe that  $p$ , if you believe each and every member of  $A$ , break down if to the beliefs in  $A$  you add those new beliefs in  $B$ . Logic and philosophy of mind thus feed each other. There is no access to the detailed workings of the mind that cannot be fixed in logical theory; no causal transitions that cannot be fit in a computational mould. For a logician, then, philosophy of mind throws the challenge of finding the way of translating causal roles into logical form relationships.

There is one more thing I wish to underlie concerning the place of logic in a computationalist conception of the mind. Logical analysis now becomes part of a project of naturalizing the mind, that is, of explaining what it is and how it works without resorting to any concept except those which belong to natural science (Fodor 1987, p.

97). Sticking to this requirement, the computationalist thinks that mental representation's logical properties and relations supervene on their physical properties, i.e., on features of neural stimulation patterns. However, this way of fitting logical properties into a naturalistic scheme forces him to confront the demand from Psychologism—Please, make clear why logical properties and relations have a normative value. This theory seems to be assumed by the computationalist in recognizing that mental computation preserves logical properties and relationships. For instance, that if you accept that it is true both that  $p \rightarrow \sim q$  and that  $q$ , then you accept that  $\sim p$  as well. Thus, since mental processes seem to respect logical properties, we are entitled to conceive human reasoning as guided by normative principles. In the end, mental norms entirely rely on the logical form of the representations involved, and form supervenes on physics.

It has been repeatedly pointed out that this naturalistic strategy is doomed to failure. On the one hand, a wide range of either mistakes of reasoning (Piatteli-Palmarini 1996; Stanovich 1999) or ways in which the theory of human reasoning differs from the theory of logical consequence (Harman 1986; 1999) has been described in the last decades. On the other hand, Computationalism views the relation between logical properties and psychological processes as extrinsic. Content, i.e., concepts and thoughts, is assigned a place in reasoning as an index to causal transitions (Loar 1981; McGinn 1989). Moreover, if we deny that its normativity is, so to speak, imposed from outside, a psychologistic explanation of logic and semantics seems to be the one possibility left (Haugeland 1997, ch. 13; McDowell 1998b; McGinn 1989, p. 190; Resnik 1985). For reasons one can perceive, e.g. in Margolis (1997, § III) and Fodor (1998, ch. 1), not everyone feels this as a threat. Finally, several attempts to naturalize content, and logical properties with it, that started off more than twenty years ago not only have not made any significant progress (Loewer 1997), but their point is still open to debate since Kripke (1982) argued that there is no objective basis for intentionality.

### *Minds as texts*

The place logical theory has in Interpretativism widely differs with the one assigned in Computationalism. By Interpretativism I mean a family of views that share the conviction that mental states and their reciprocal transitions make up a kind of text or interpretable stuff. The key to this text is provided by the speaker's behaviour, both linguistic and non-linguistic, a stuff which is put under a triangulation technique that looks for the best fit among the speaker's beliefs and desires, the meaning of her language, and the way the natural and social world she inhabits and acts upon is. This way of interweaving mind, language and world naturally paves the way to the idea that mind is recognizable where there is room for interpretation. This is so because the basic subject matter of interpretation is intentional behaviour, action; not naked movement, but movement filled with sense; movement successfully placed in a web of beliefs, intention, expectances, desires, etc. The need to make sense of action requires that the interpretee's mental states and intentional actions cohere not only among themselves but to each other as well. Coherence is rationality, i. e. the kind of fit that makes both thought and action satisfy justification constraints.

Among the many guises that Interpretativism adopts, two have circulated widely. In Donald Davidson's approach the intentional agent has to be radically interpreted. This requires that the interpreter has to make himself familiar with the native's language and culture from scratch. He knows nothing of her community's language, nothing of her systems of beliefs and values (Davidson 1980; 1984). A hypothetical point of departure like this neatly exhibits the reciprocal dependence of language and thought. There is no route to what the native's utterances mean in her community that does not cut across her beliefs and desires. And conversely, no route to most of her thoughts and their contents that does not need an authoritative knowledge of her language. With those restrictions radical interpretation starts off and goes forward, provided that the interpreter assumes that he is having dealings with a community of rational agents, men and women that believe what is manifestly true and desire what is manifestly good. The articulation of this intuition has given rise to the so-called Principle of Charity (Davidson 1984). It guarantees that an intentional agent's beliefs and desires, as well as her language, are linked to each other and to the world as they should. In this variant of Interpretativism, assigning content to mental states and meaning to words and sentences is a product of the best balance between forces that, coming from the world, mould the agent's natural condition, inducing on him perceptions, beliefs and other mental states, and forces that bestow coherence on those multiplicity of states and reciprocal transitions. Intentionality is a result both of the agent's natural condition and of the network of those beliefs and desires he comes to acquire.

According to the variety of Interpretativism sponsored by Daniel Dennett, the intentionality of language and thought lies in the interpreter's eye. All intentionality is derived. It so happens that there are behaviours and other events that are only made sense by adopting what Dennett has called the intentional stance, the kind of attitude that prompt us to treat things as if they had the sort of organization characteristic of intentional agents. By so acting we predict and explain what they will do or have done, by attributing them desires and beliefs. An intentional system—an electronic chess player, a lioness stalking her prey or a human being—is anything whose behaviour makes sense when seen from such a stance (Dennett 1978; 1987). Not anything is an intentional system. Thus, only in extraordinary situation carburettors, trees and jellyfish are given this title. There does not seem to be a neat answer to the question of where lies the border between intentional and non-intentional systems.

The way logic integrates into the interpretativist view of mind is dictated by the Principle of Charity. Logical theory substantially contributes to this view by making explicit what conditions of consistency and logical consequence a system of cognitive and conative states has to abide by if it has to be a rational one. Thus, a certain amount of logical theory has to be credited to the native on pain of not being conceived by the interpreter as a bearer of manifest truths and values. How much—that has been open to a variety of opinions. As far as logical theory is concerned, Quine (1960, ch. 2) maintained that the speaker's logic does not go beyond the basic principles that govern the propositional connectives ('not', 'and', 'or'). Davidson (1984) extended to classical quantification the amount of logic that the interpreter has to see reflected in the native's thought and linguistic abilities theory. Though these are two

ways of substantiating the Principle of Charity *vis-à-vis* logical theory, other manoeuvres are still possible. In fact, it seems reasonable to think that the specific content the Principle may have would depend on what the interpreter's constraints on logical consistency and consequence are.

Now, it is an interesting feature of Interpretativism that, given the centrality it assigns to the Principle of Charity, this conception of the mind easily accounts for the normativity both of content and meaning. The intentionality of thought and language is of a piece with the rationality of mental states and intentional action, the reason being that there is content and meaning only where norms of justification concerning logical coherence are in force.

The last point helps one to become aware of an important difference between Computationalism and Interpretativism, namely, their different attitude towards the doctrines according to which the basic properties of the world are physical. While Computationalism's vocation is openly naturalistic, Interpretativism in some of its variants rejects this creed. In Davidson's view, this is a tribute we have to pay for the "fundamental difference [there is] between how we come to know what others mean and think and how we come to know, if we do, how the blood circulates or planets are formed" (Davidson 1994, p. 126). As for computationalists, such a difference is illusory. Intentional properties are properties had by certain inhabitants of the world, and "if the ontological possibilities aren't epistemologically constrained in physics, why should the ontological possibilities be epistemologically constrained in linguistics or psychology?" (Fodor & LePore 1994, p. 102). On the one hand, his rhetoric question implicitly carries with it the denial that both content and meaning are relative to patterns and practices of interpretation. On the other hand, however, in doing so Computationalism at least implicitly claims that the normativity of logic is not fact worth explaining.

### *From minds to persons*

While Computationalism and Interpretativism have for the last decades commanded a lot of the attention of the philosophical world, many philosophers have considered their own views distant from each of those. As early as in the forties and fifties Gilbert Ryle and Ludwig Wittgenstein created a somewhat blurred tradition of thought that have cleared up other paths. Ryle advanced an influential argument whose target were not only Cartesian dualists, but their physicalist opponents as well. According to it, both Cartesian Dualism and Physicalism err in wondering what stuff, whether matter or mind, we are made of. In Ryle's view (Ryle 1949), this question already contains the seeds of confusion, because to think and talk about anyone's mind is to think about her abilities, to think and talk in semi-hypothetical terms about people's behaviour. Peter Geach (1957/1992) showed that Ryle's dispositional view of the mind is incomplete, because many mental states do not have behaviours characteristic of them. The idea that perceptions, concepts, judgements, desires or emotions might belong to an independent ontological category, one that requires a pattern of conceptualisation made to measure, began to break through and quickly settled down among many phi-

losophers after Peter Strawson's (1959, ch. 3) suggestion that all of them are properties of persons, and persons are not either immaterial minds or material bodies.

A way of articulating the intuition just mentioned takes desires, beliefs, perceptions or emotions as part and parcel of causal psychological explanation, and insists on how far is this kind of explanation from the one used in physics or chemistry (Baker 1995). Another view (Wittgenstein 1953/2000; Sellars 1956/1997; Hornsby 1998; McDowell 1994; 1998a) has it that, in addition to conceive reasons as full causes, understanding human mind demands that its states and events be placed in a space of reasons. To understand is not to exculpate but to justify, as McDowell (1994) has put it. Following Hornsby, I will talk of Naive Naturalism to refer to this conception of the mind, a label inspired in the belief that it is possible to delineate a non-scientific view of person's nature. (This proposal, either present or found in Wittgenstein by many of his readers, connects the ideographic tradition of continental philosophy to English language Analytic Philosophy.)

The question now is what role logic has been thought to fulfil in this conception of the mind. From an interpretative point of view, behaviour, linguistic or not, has to respect logical principles in order to be intelligible. Any event admits of more than one description, but it emerges as rational when it is given an intentional description. Naive Naturalism has it that every description of a mental state of theirs is subject to rationality demands, to norms of justification. The constitution of these norms ties down the relationships among mental states themselves and among mental states and states of the world to conditions that would not be met were those relationships conceived as effects of neural mechanisms. In particular, no light may be shed on logical abilities if, in what would be a vindication of Psychologism, it is claimed that they are causal effects of those mechanisms.

In recent decades both Michael Dummett y John McDowell have insisted on this criticism, though they have disagree on the means to block the way to Psychologism (Dummett 1973/1991; 1993; McDowell 1998b). The problem of what format a theory of meaning should have has been the arena on which Dummett and McDowell have confronted their views on meaning and mind. They agree that a theory of meaning has to fix the contents of speech acts, both effective and possible, performed in a linguistic community. In so far as a member of a community has command of the language used in it, a theory of meaning that assigns her speech acts the right content may be thought of as systematizing the speaker's understanding of her language. The degree of fitness of the theory can be measured by taking into account how acceptable the descriptions of speaker utterances that the theory provides are. Those descriptions must display speaker utterances as performances of speech acts with a specific content, thus making them understandable in virtue of their relations to the speaker's desires, beliefs and further mental states.

How a theory of meaning explains that capacity is the point at stake between Dummett and McDowell. In Dummett's view, the theory has to be robust: for each semantic value that it assigns to an expression of the language, it must contain a clause that somewhat states the public conditions in which any speaker of the language would recognize that the expression is adequately used. In addition, the theory has to

be framed in a conceptual system that does not presuppose what it should explain. This implies that the theory is forbidden the use of terms like ‘...expresses the thought that...’, ‘...states that...’ or ‘asks whether...’. It is because it abides by this constraint that a meaning theory is immune to Psychologism. Moreover, from the assumption that a theory of meaning indirectly deals with a rational capacity, Dummett gathers that it is also free from the dangers of behaviourism. On the contrary, McDowell thinks that a theory of meaning has to be modest: it has to map language expressions onto semantic values on the condition that the mapping helps making sense of both the speaker linguistic behaviour and her linguistic partners’ reactions. And making sense of it is something the theory will only achieve if it manages to adequately describe the speaker utterances as performances of speech acts in a medium of mental states subjected to rationality constraints. (Thus, a non-intentional description of the sort of mental abilities involved in language use, i.e., a description of a complex neurobiological mechanism, is of no value for a meaning theoretician.) As a consequence, Dummett’s manoeuvre against Psychologism is inefficient because “in theorizing about the relation of our language to the world, we must start in the middle, already equipped with command of a language” (McDowell 1998b, p. 330). We are not cosmic exiles. There is for us no mean of analyzing meaning and content with the help of conceptual materials from which our prior involvement in characteristically human forms of life has been removed.

The question whether a theory of meaning should be robust or modest has deep implications on the debate on Psychologism and the normativity of logical principles. The question is, is the ability to deploy these principles a constitutive part of the speaker’s faculty of understanding? Dummett believes that the theory of meaning has to be robust: any the presence in the theory of any principle of that gives the meaning of a logical connective (‘not’, ‘and’, ‘or’) should be under the control of a clause that puts forward the conditions under which the speaker becomes aware that the application of the principle is justified. The normativity of each logical principle thus gets its own justification. However, he concludes, this constraint makes us suspicious of the Law of Bivalence—every sentence is either true or false—and should lead us to choosing intuitionistic logic as base of a theory of meaning for a natural language. As contrasted with Dummett’s, a modest view is less demanding. The normativity of logical principles comes as a consequence of their being part to a theory of meaning that is globally checked for its success in accounting for the speakers’ capacities in thought and linguistic communication, and no separate justification of any logical principle is required. The real contribution made by the logical apparatus of a semantic theory does not go beyond guaranteeing that the meaning of a complex sentence depends on the meaning of its sentence constituents. Because of it, principles like the usually yielded in truth-tables (‘ $\neg a$ ’ is true iff ‘ $a$ ’ is not true) are separately subject to norms of justification. What logic is the right logic is not a problem for the theory of meaning.

### Conclusion

Three conceptions of the mind and of how they make sense of logic have been put forward. For Computationalism intelligence, and rationality with it, is a mechanical affair, and logic theory provides an abstract set of transitions among machine states. Interpretativism sees the attribution of logical abilities to speakers as following from the demands of the Principle of Charity as a condition for the possibility of radical interpretation. Naive Naturalism sees logic as a precipitate of rational abilities that underlie language use and understanding.

### BIBLIOGRAPHY

- Baker, L. R. (1995). *Explaining Attitudes*, Cambridge University Press.
- Davidson, D. (1980). *Essays on Actions and Events*, Oxford University Press.
- Davidson, D. (1984). *Essays on Truth and Interpretation*, Oxford University Press.
- Davidson, D.: 1994: "Radical Interpretation Interpreted", *Philosophical Perspectives* 8, 121-128.
- Dennett, D. (1978). *Brainstorms*, Vermont: Bradford Books.
- Dennett, D. (1987). *The Intentional Stance*, M.I.T. Press.
- Dummett, M. (1973/1991) *Frege. Philosophy of Language*, London: Duckworth.
- Dummett, M. (1993). *Origins of Analytical Philosophy*, London: Duckworth.
- Fodor J. (1980). "Methodological Solipsism as a Research Strategy in Cognitive Psychology", *The Behavioral and Brain Sciences* 3, 63-109.
- Fodor, J. (1987). *Psychosemantics*, M.I.T. Press.
- Fodor, J. (1994) *The Elm and the Expert*, M.I.T. Press.
- Fodor, J. & LePore, E. (1994). "Is Radical Interpretation Possible?", *Philosophical Perspectives*, 8, 101-119.
- Fodor, J. (1998). *In Critical Condition*, M.I.T. Press
- Geach, P. (1957/1992). *Mental Acts*, Bristol: Thoemmes Books.
- Harman, G. (1986). *Change in View*, M.I.T.Press.
- Harman, G. (1999). *Reasoning, Meaning and Mind*, Oxford University Press.
- Haugeland, J. (1998). *Having Thought*, Harvard University Press.
- Hornsby, J. (1997). *Simple Mindedness*, Harvard University Press.
- Kripke, S. (1982). *Wittgenstein on Rule and Private Language*, Oxford: Basil Blackwell.
- Loar, B. (1981). *Mind and Meaning*, Cambridge University Press.
- Loewer, B. (1997). "A Guide to Naturalizing Semantics", in Hale, B. & Wright, C., eds.: 1997.
- Margolis, J. (1997). "Late Forms of Psychologism and Antipsychologism", *Philosophy & Rhetoric* 30, 291-311.
- McDowell, J. (1994). *Mind and World*, Harvard University Press.
- McDowell, J. (1998a). *Mind, Value, and Reality*, Harvard University Press.
- McDowell, J. (1998b). *Meaning, Knowledge, and Reality*, Cambridge: Harvard University Press.
- McGinn, C. (1989). *Mental Content*, Oxford: Basil Blackwell.
- Piattelli-Palmarini, M. (1996). *Inevitable Illusions. How Mistakes Rule Our Minds*, London: John Wiley & Sons.
- Quine, W. (1960). *Word and Object*, M.I.T. Press.
- Resnik, M. (1985). "Logic, Normative or Descriptive", *Philosophy of Science* 52, 221-238.
- Ryle, G. (1949/2000). *The Concept of Mind*, London: Penguin Classics.
- Sellars, W. (1956/1997). *Empiricism and the Philosophy of Mind*, Harvard University Press.
- Stanovich, K. (1999). *Who is Rational? Studies of Individual Differences*, Mahwah: Erlbaum.
- Strawson, P. (1959). *Individuals*, London: Methuen.
- Turing, A.. (1937). "On Computable Numbers, with an Application to the Entscheidungsproblem" in *Collected Works of A. M. Turing: Mathematical Logic*, R. Gandy, C. Yates (eds), Amsterdam: North-Holland, 2001.
- Wittgenstein, L. (1953/2002). *Philosophical Investigations. The German Text, With a Revised English Translation*, Oxford: Blackwell Publishing.



**Juan José ACERO** is professor at University of Granada and Director of Department of Philosophy at such University. He is specialized in philosophy of language, philosophy of mind and history of analytical philosophy, areas in which he has published three books and around fifty works counting articles and chapter of books. His present-day themes of work are the following: nonsense, intention and emotion. He is editor of *Teorema* (1998) that commemorates the twenty-five years from publication of "Naming and Necessity" and of the volume about *Filosofía del Lenguaje. I Semántica* of the *Enciclopedia Iberoamericana de Filosofía* (1998).

**Address:** Depto. de Filosofía, Facultad de Filosofía, Universidad de Granada, Campus de la Cartuja, 18011 Granada. E-mail: [acero@platon.ugr.es](mailto:acero@platon.ugr.es)