A Fodor's Guide to Cognitive Science

## ROBERTO G. DE ALMEIDA

The so-called cognitive revolution-the second, by some accounts, after Descartes'-began taking shape over 60 years ago. Intellectual revolutions, as you probably know, are almost never the labor of a single mind (perhaps, again, sauf Descartes'). They are usually the result of scientific and philosophical discontent with modes of explanation and with the very nature of the explanandum. And they lead inexorably to changes in theory and empirical object. Or so they should. The "second" revolution on the workings of the mind brought forth a torrent of new guiding assumptions in linguistics, psychology, computer science, and philosophy, among other disciplines. In this revolution, heads did not roll: they turned. The history of these disciplines and how they came to be together under the big tent of cognitive science cannot be reduced to just a few names, even if they are the names of true pioneers. But history is unfair and sometimes reduction is the only feasible way to convey the transformations a field might go through: the proper names become metonymic for the ideas, the ideas become standard (or, as it happens, generate classical controversies), and the history of the revolution is largely told by the names of those who are taken to push it forward. The short history I want to tell is like that. No matter how one maps it out, it has Alan Turing as an early influence, even if his influence was felt only later. And, of course, Noam Chomsky and modern-Chomskyanlinguistics are mainstays. Along the same lineage, the cognitive revolution owes Jerry Fodor some of its most fundamental ideas. Perhaps this lineage should be traced back to Plato, Ockham, Descartes, Locke, Hume, and a few others, with no clear discontinuity-certainly passing by Frege, the early Russell, and the early Wittgenstein. But in contemporary work, Fodor's name is metonymic with a kind of cognitive science—possibly *the cognitive science*—that many of us care about doing. One could refer to it as *Fodorian* cognitive science.

The chapters collected in this volume are a celebration of that kind of cognitive science, of its most fundamental ideas, and, in particular, of Fodor's contributions to psycholinguistics, to the theory of concepts (thus, to a theory of the elements of the *language of thought*), and to cognitive architecture, more broadly. I won't really call it *Fodorian* cognitive science because Fodor's contributions are so entrenched—and so inspiring and, because of that, at times so controversial—that I like to call it simply *cognitive science*. But this volume does not constitute the kind of celebration you might expect, the homage he would refuse. In line with his polemical style, the goal was to bring to fore a critical evaluation of the foundations of many of these ideas; we<sup>1</sup> wanted to put them to test, but also to move them forward (or, if it's the case, move away from them). We wanted, in sum, to examine the status of these ideas and how they might set the agenda for what is to come.

Now, here is some background on how we got thus far. Fodor's main contributions to cognitive science gather around language and thought and, of course, the nature of the language *of* thought, its elements—concepts—and how they connect to the world. This is not to say that all his philosophy of mind and philosophy of language—not to mention his many incursions into experimental work—can be *reduced* to language and thought, but these are the key terms of a deep and wide theorizing about the nature of the human mind.

Fodor entered the scene around 1959 when, coming out of doctoral work with Hilary Putnam in Princeton, he arrived at MIT. The "second" cognitive revolution, then, was "in the air." Chomsky was there, himself surrounded by behaviorist territory-and at striking distance. George Miller was then infiltrated in that territory, at Harvard, starting a program of research that was full-blown cognitivist, having among his aims the marriage between the new linguistics and a psychology that was increasingly soaking up computational metaphors.<sup>2</sup> The story is long and its plot reads somewhat like a thriller, as far as intellectual thrillers go. Chomsky (1959) had just famously exposed the limits—and the explanatory inadequacies—of behaviorism: to put it simply, there had to be a mechanism underlying both language attainment and language use, and that mechanism was far more complex than simply pairings of stimuli and overt or covert responses. The plot thickens, for much of this revolution was also taking place elsewhere—in computer science and in psychology, in domains such as in memory and attention. But here the focus will be mostly on the bits of how psycholinguistics and the core of cognitive science became what it is today (or what it is supposed to be).

At MIT, in the early 1960s, Fodor (re)encounters other young co-conspirators, including Tom Bever and Merrill Garrett, both of whom had been infiltrated in behaviorist territory before: Bever at Harvard, Garrett at the University of Illinois. Fodor had been a visitor at Illinois, where he exchanged ideas on the nature of psychological explanation with Charles Osgood, one of the leading behaviorists then. I put this very politely, because this visit—and those

exchanges—only deepened the canyon that separated the two then competing worldviews, particularly on what pertains to language and cognitive processes.<sup>3</sup> Osgood later remarked that science can be "faddish,"<sup>4</sup> with which I agree, but for reasons that will soon become clear.

Meeting his younger co-conspirators was instrumental in advancing the cognitivist brand of psycholinguistics into enemy territory. It was then and thence that, with the little camouflage that Jerry's old Austin Healey provided while crossing Cambridge, MA, or in the trenches of their improvised lab, Jerry, Tom, and Merrill plotted about changing psycholinguistics; or perhaps plotted about new ways of testing linguistic postulates deploying experimental methods. The psycholinguistics of the early 1960s was still dominated by what we can call "psychology of language," mostly destitute of its linguistic core. There were, of course, very notable exceptions: Miller was then the main driving force behind a new linguistically informed experimental psycholinguistics. What was in the air—or in a few of those minds-helped establish the materialistic mentalism that was rejected by the dominant "behavioral science." Crucial to these advances was Chomskyan linguistics (the adjective stands for what was then already a small legion), which was beginning to thrive, thus providing the impoverished psychology of language with the algorithms it was supposed to run. To be clear, that's not the beginning of psycholinguistics, for since the early 1950s the term was already being thrown around, labeling other forms of contact between the structuralist linguistics of thence, experimental psychology, and theory of communication. And certainly that was not the beginning of experimental psychology of language, which can be traced back to James Cattell and Wilhelm Wundt. But it was the pinnacle of theoretical work on the formal properties of the apparatus that yields a language—and, by hypothesis, the mechanisms for its use-combined with the experimental paraphernalia of incipient cognitive psychology, both heavily guarded by philosophical functionalism about psychological explanation. Those were the beginnings of Cartesian Psycholinguistics.

A small portrait of Turing could have been hanging there somewhere, perhaps in the improvised lab, as a reminder of the agenda for cognitive science (nobody called it that, then): the prospects for a theory of language hinged on understanding the nature of its underlying rules and representations. And so did the prospects for a theory of mind (at least some of its main faculties or modules, as we'll see). Many experiments ensued and many techniques were developed, beginning with the "clicks" that perceptually (and illusorily) marked the boundaries between clauses within sentences. We were then beginning to "see" what the mind does when a sentence comes its way. The black box was cracked open; rats and pigeons were spotted in the unemployment line. The results of this collaboration appeared in many experimental and theoretical papers spanning over 10 years, with its apotheosis being Fodor, Bever, and Garrett's classic The Psychology of Language: An Introduction to Psycholinguistics and Generative Grammar (1974).<sup>5</sup> The one-liner would read like this: the computations involved in understanding a sentence do not go pari passu with the transformations that grammatical principles determine for its structural analysis-from its surface

form to its kernel—nor are the computations effected by some analysis by synthesis in which the grammar provides a set of possibilities, a search space. The computations involved in sentence comprehension rather rely on heuristic procedures, perceptual strategies for analyzing sentence constituents based primarily on the detection of clause boundaries together with analysis of constituent structure within those clauses.

There were at least two main arguments for not tying the computations involved in sentence perception to actual sequences of syntactic transformations; and these arguments are as valid today as when they were first put forth, even if both grammatical theory and parsing models have moved away from transformations. One is that the putative linear perception of a sentence allows, at every moment, for myriad possible structures compatible with the input, yielding a "search space" that is less than practical, perhaps close to impossible, for quick online structuring and interpretation. Another is that different transformations (when we take movement of constituents into account) lead to partial analyses that can be deemed incompatible with surface input—even ungrammatical—raising the need for almost constant course corrections in parsing and interpretation.

Phrase structure grammar, it should be clear, was said to underlie sentence structure: it was "psychologically real," as people used to say back then. That was not under dispute within the cognitivist camp. But the process by which phrase structure is perceptually computed was said to rely on independent phrasal packaging mechanisms. This view later evolved into different parsing models, from the "sausage machine," proposed by Lyn Frazier and Janet Dean Fodor, all the way to the "good enough" approach that Fernanda Ferreira has championed more recently.<sup>6</sup> Parsing models have since swung between these views—with linguistic principles either actively engaged in the moment-by-moment analysis or operating on the product of other processing routines that are taken to be more or less dedicated to processing language (I will leave aside those who believe there are no linguistic principles at all).

Even if the empirical data did not fully support the clause as the perceptual unit, the 1960s and 1970s psycholinguistics chapters of the cognitive revolution became paradigmatic of what cognitive science came to be (or was supposed to be): largely interdisciplinary, a collaborative enterprise without boundaries between established disciplines or departments (well, now I may be daydreaming). It is not the case though that all rebels were speaking in unison but their voices were then shouting in a similar direction. The beauty of rationalist revolutions is that no blood spills on the streets.

A few years before, in his *Psychological Explanation* (1968), Fodor had focused on the metatheory for cognitive science—or at least on one of its main philosophical approaches: *functionalism*. Along the lines of Chomsky's (1959) attack on B. F. Skinner's *Verbal Behavior*, in *Psychological Explanation*, comes a more detailed plan for attacking the philosophical foundations of behaviorism and its positivistic roots—but chiefly the anti-mentalism represented mainly by G. Ryle. The emphasis was on understanding what constitutes an explanation

in a "special science" like psychology. The plan had two fronts. One was the idea that an explanation of "behavior" could not dispense with the underlying causes of which overt behavior is only a consequence (and limited at that). Another was the rejection of reductionism for psychology. Vienna gave us great music (as Fodor knows it well) and great philosophy (though he might question its consequences), but the logical-positivistic thesis that eventually all sciences could be reduced to physics did not bode well for psychology, not at least for *cognitive* psychology.

Both in Psychological Explanation and in papers collected later in Representations: Philosophical Essays on the Foundations of Cognitive Science (1981), Fodor argued for the special status of functionalistic explanations. If you are in doubt, here is a clarification: functionalism in philosophy of mind is materialism, but the kind of materialism that does not appeal to the nuts and bolts of the machine (or neurotransmitters and neuroanatomy, for that matter): it takes functional properties to suffice at a certain level of explanation. This level is something Alan Newell, Zenon Pylyshyn, and other bona fide cognitive scientists have called "symbolic" or what David Marr called the "representation and algorithm" level. Whichever label one chooses for it, or whichever way one partitions the analysis, it is at the symbolic/algorithmic level that a cognitivist explanation about rule-governed processes is best conceived. And it is also perhaps at that level where we should begin to approach the so-called knowledge-based processes, the intentional kinds. (Not to be forgotten: cognitivist/functionalist explanations also appeal to folk-psychological mental states or attitudes: it is because I planned to write these very words-following a long chain of desires, beliefs, hopes, doubts, and (in)actions-that I actually came to do it.) And no matter how one conceives of this relation-between the symbols and what they represent, between the rules and their following in the course of mental processes-no particular status is given to the "other" level, the biological or implementational one. The issue is about explanation and qua explanatory level, appealing just to biology won't do. Of course, one should not be in any way discouraged from actually pursuing an investigation of the biological level, quite the opposite. Functionalism is materialism, one must insist: It is assumed that functional processes supervene upon physical ones, as Fodor keeps saying. But before the functional magnetic resonance imaging (fMRI) machine is plugged in, one needs to make sure to have a good working theory at the symbolic/algorithmic level.7 Revolutions are more effective when transformations take place at the infrastructure; and, in the cognitive revolution, the infrastructure is not in the biological bases of behavior but in the functional mechanisms that the biological substrate executes. There is, however, a very active minority of cognitive neuroscientists (the prefix is more fashionable, these days) who are realists about representations—who follow something like a methodological law: neuroscientists who are cognitive scientists have to postulate representations.8

There is for sure a direct connection between the battles fought in psycholinguistics and in philosophy of mind. The cognitive revolution needed to have its

guiding assumptions about the object of investigation clear: the internal states of the organism. But at the same time it also needed to show that its theoretical and empirical accounts of these internal states had validity. This amounted to both, postulating the nature of the internal system of representations and processes underlying those states and providing empirical evidence for their workings. Psycholinguistics was at the center of this program of research because it had all the ingredients necessary to build a theory of mind: it had the symbols and algorithms for linguistic computations, and a mechanism for yielding internal representations. And in true cognitive-science fashion, no type of evidence was ruled out. Just as distributional arguments and crosslinguistic evidence were important for advancing *linguistic* arguments, they were also important for advancing arguments about the nature of the mind's *internal code*. Experimental evidence—coming from all corners of booming cognitive psychology—was also instrumental in pushing forward the agenda. Fodor has been committed to the science of the mind and to its philosophical foundations, and very few-in the last couple of centuries-have been able to keep these commitments as prolifically as he has; and few have transited between the science and the philosophy of mind with the same ease.

The Language of Thought (LoT, 1975) is typical of this attitude, a landmark for psycholinguistics and for the view that the mind is best conceived as a computational device. It is in LoT where we see three of Fodor's main threads coming together like in no work before: philosophy of mind/science, psycholinguistics, and the roots of his Computational/Representational Theory of Mind (henceforth, C/RTM). It is in fact his view of cognitive architecture that begins to emerge, with implications for several of his lines of work. LoT was much more than the "speculative psychology" he claimed it was. It detailed what a commitment to C/RTM entails:<sup>9</sup> first there ought to be representations, if explanations appeal to anything other than simple overt behavior, and representations are the medium for processes, which are carried out as computations. LoT had it moreover that many mental states were relations to propositions, that to believe or to desire *P* was to be in a relation—a *computational* relation—to a representation of P, which was couched in the vocabulary of the internal code. Computations leading to mental states were taken to be sequences of events akin to derivations (e.g., the sequences of syntactic operations; or the sequences from premises to conclusion in syllogistic reasoning). This is in essence what constitutes the common operations of putative cognitive processes. And, by hypothesis, the language of thought bears many of the properties of natural language: it is recursive, productive, compositional, and it is a typical computational system, for its processes too are computations over (symbolic) representations. There is a caveat, though: as Fodor warns us in the last chapter of *LoT*, quite possibly a few ("more than none") cognitive processes behave that way, but most likely not all do. Cognition is to a large extent holistic, context-sensitive (think about, e.g., decision making). And there might be lots of propositional attitudes that are not computationally derived—for example, those whose causes are not psychological.<sup>10</sup> But if we were to have a (cognitive) psychology, a good way to start was to devise a theory of

the internal representations and how these representations were manipulated in mental processes.<sup>11</sup>

The plan for cognitive science taking over all (relevant) psychological accounts **L** of typically *cognitive* processes was not complete, of course. First, because there was no detailed plan to follow: cognitive science from its inception has been anarchic, and it was then barely holding on to a few postulates on what constitutes the proper level of analysis. And second, because the conception of the mind that was then emerging begged many questions: What was the nature of the code? Or how many codes were there? Which processes were supposed to be computational and which ones were not? As more specific hypotheses about the nature of representations and processes were ironed out, yet deeper questions internal to the program were raised. One of Fodor's key concerns was mental content-roughly how symbols get to represent what they do and how they enter into putative intentional processes. This appears early on in LoT and in the original essays of Representations. In fact, accounting for the nature of the units of representation—let's call them *concepts*—became one of Fodor's main missions, spanning over 50 years of hard labor. And not surprisingly, this is perhaps the central issue in cognitive science, for it underlies many others, from the nature of visual processes of object recognition, to language comprehension and production, and certainly to many "high-level" processes we can call thinking. If concepts are the building blocks of the representations manipulated in all these processes, if they are the building blocks of all processes that employ anything having to do with content (all that's relevant about perception and cognition, as far as I can tell), then how are they represented, and how are they developed in the organism?

Fodor once said that every Monday morning there was a meeting at MIT to decide what would be innate that week; whoever had the most outrageous proposal would chair the works. I don't think this is entirely a joke, as it is clear that nativism of some sort is the only route to the postulation that internal states develop and change partly in response to environmental causes. Poverty of stimulus arguments stand not only for language but for concepts too. It was in this context—perhaps in one of those Monday morning meetings—that conceptual nativism became central to Fodor's work. In his early treatment of conceptual nativism, he showed that the process of concept attainment couldn't be anything nearly what many cognitivists and practically all empiricists were postulating it was: a process of learning. More than an assertion, there was an argument, a puzzling one. Fodor suggested that what was being shown by Jerome Bruner, Jean Piaget, and others, under the rubric of "concept learning," was what he called "belief fixation." Roughly, decisions about the extension of a given word/ concept—say, wyz—presuppose the existence of the criteria (features or properties such as ROUND and GREEN) upon which those decisions are based. Thus, what the organism has at its disposal are the very premises for inductively fixating the belief or hypothesis that the referent is a WYZ (and all this requires, of course, a vocabulary of representations, a language of thought). This kind of argument made strong waves in the canals of the Abbaye de Royaumont, near Paris, where, in 1975, Massimo Piattelli-Palmarini brought together Chomsky, Fodor, and other nativists for an epic debate with Piaget and his constructivist colleagues. Legend has it that some of the best arguments pro nativism still echo in the Cloister.<sup>12</sup>

Nativism about concepts, contrary to the popular joke-that we know the likes of AIRPLANE and ELECTRON from birth-assumes that the conceptual stock must be primitive. The problem is that, on pain of committing to analyticity (see later discussion) or, worse, to the idea that concepts are structured (the problems do overlap), the conceptual stock has to be vast, having more than just the sensory primitives of classical empiricism. Even the classical empiricists—Locke and Hume—were committed to some form of nativism, except that their commitment was to the sensory basis or to the conditions for picking out the sensory basis. But the sensory apparatus-or what the sensory apparatus, by hypothesis, yields—underdetermines the bases upon which belief fixation relies. Hence, the only way out of this morass is to assume that indeed the conditions for fixating AIRPLANE and ELECTRON are innate. It's the structure of the mind that allows for the triggering of concepts by experience. And because all concepts are acquired like that or because most concepts are triggered like that, they ought to be considered all primitive, atomic, not molecular.

As the reader surely noticed, despite all denials, Fodor flirts with empiricist postulates, but not with the kind of empiricism that is radically anti-nativist. In fact, he denounces a strict dichotomy between empiricism and nativism. Fodor is empiricist with regard to the primacy of the perceptual input in causally determining—or triggering—the conceptual stock. Since all lexical concepts are primitive, or all lexical concepts arise from primitive functions, the main worry is how the organism works on triggering or fixating its supposedly vast stock. He assumes that it is probably the basic level—*DOG*, not *POODLE* or *ANIMAL*— what is first triggered by the environment, and one works out different levels of generalization or specificity along the way. Notice that, contrary to what one would suppose—if classical empiricism were to be enforced—it is not *RED* and *LINE* that the child picks up, but putative links with referents that are possibly at the basic level of abstraction. And even in the case of *RED* and *LINE*, what determines their primitive status is not that they are sensory, it is that they bear *properties*.

The arguments in *Representations* and *LoT* surely raise lots of questions. The work of understanding how concepts get linked to their referents requires fine-tuning a cognitive architecture that affords these links. Enters, then, *The Modularity of Mind* (1983). Though *Modularity* does not appear to be "causally connected" to the early work on concepts, it plays an important role in Fodor's program. It is the centerpiece of much of his work linking the psychology of perception, C/RTM, and the idea that higher cognitive processes involve large doses of belief fixation. It is *via perception* in fact that belief fixation begins to take place, with the triggering or matching of concepts by referents. Perception—at

least the classical empiricism way—was a process of matching a thing to an Idea, a process that was atomistic, as Fodor noticed. It is somewhere here that the modularity of perception and atomism about meaning meet: roughly, seeing a cow triggers *COW* (even if you think there are features, seeing a spot or a horn triggers *SPOT* or *HORN*).

The story about atomism in conceptual representation and the story about the modularity of perception, then, are complementary: if you believe, as Fodor does, that much of conceptual tokening is "brute-force" linking between referents and their representations, you are somewhat committed, as he is, to the modularity of perception. What you see-to stick to vision-is independent of what you believe; and the concept you token is too independent of other sorts of beliefs you might have. What Fodor proposed in Modularity, more specifically, is that the perceptual analysis process is highly constrained. In his version of modularity, Fodor takes perceptual analysis to be encapsulated from the rest of cognition, with modules, notably vision and language, separate from each other and from other systems. The modules have their own rules and have access to their own representations, mostly the ones that are causally connected to the analysis of input post-transduction-they are causally connected, in sum, to the kinds of stimuli that are the modules' natural kinds. Crucially, modules, in their task of producing perceptual analyses, are not influenced by the beliefs that the organism has at its disposal. It is here where Fodor traces the line between perceptual computations and what he called the holistic, Quinean, central system, where all outputs of modules eventually meet. There is an epistemological thesis here as well: observation and inference ought to be kept apart, just like perceptual computations and beliefs ought too.

I won't say much more about modularity because several of the chapters in the present volume assess the modularity hypothesis, what became of it, and even how it can be reframed in current cognitive science.<sup>13</sup> But I want to call attention to Modularity's sizeable impact in the psychology of perception, where it set the agenda, the guiding hypotheses on how language (in particular, but not exclusively) might be perceived. Fodor's formulation assumes that the module for language is dedicated to input analysis (though here he spars with Chomsky),<sup>14</sup> producing what ought to be minimally some form of syntactic or perhaps something like a logical representation of the linguistic input. The general idea of modularity was as well in the air when Fodor wrote his influential monograph, but he refined the hypotheses and marked the boundaries between two main psycholinguistic camps: those who assume some level of autonomy for language perception (and its internal computations) and those who assume perception to be, in the term coined by Pylyshyn, "cognitively penetrable." Most parsing models from the early 1980s were predicated on how much or at what point in time they allowed for non-linguistic information (non-sentential context, beliefs, expectations) to influence structural decisions. This issue has never really been settled. And although I am not keen on appealing to arguments from philosophy (or sociology) of science to legislate on matters in need of theoretical

and empirical treatment, it is worth emphasizing that, as Feyerabend (1975) once put it,

No idea is ever examined in all its ramifications and no view is ever given all the chances it deserves. Theories are abandoned and superseded by more fashionable accounts long before they have had an opportunity to show their virtues. (p. 35)

Maybe the modularity hypothesis is not at that stage yet—it has neither been abandoned nor superseded, despite the enormous amount of research conducted on behalf of its constituent postulates. But it is clear that fashions change—and research grants go with them.

We have to admit, then, just like in Osgood's reaction to the new psycholinguistics in the 1960s and 1970s, that science can be faddish. But it seems that, in its current stage, cognitive science does not have many viable alternatives other than to assume—as a working hypothesis—that some of its main systems might be encapsulated and, moreover, that some or perhaps most of its representations and processes are symbolic and computational. One might think, of course, of scores of alternatives to the architecture that C/RTM breeds. Think for instance of connectionism, which was trumpeted, when it came out in the late-1970s, as a revolution (within the revolution, I suppose). Connectionism was supposed to provide cognitive science with what it appeared to lack: some strong neurological plausibility; it was supposed to rescue physicalism while holding on to the idea that representational states (the activated nodes) are entertained in the course of cognitive processes. Moreover, what gave connectionism its most plausible selling point was the idea that representations were causally connected as if they were (actual) neuronal networks-with their activation and inhibition functions as well as learning capabilities operating as massively parallel, interconnected units. But soon-perhaps not soon enough-it became clear that connectionism failed to account for many of the key properties that C/RTM took to be front-and-center. Fodor's move to Rutgers University, in the late 1980s, afforded a closer collaboration with Pylyshyn, at the Rutgers Center for Cognitive Science, which they founded (not to be discounted were also the strategic proximity with the opera at the Met and the sailing on the Hudson). They were then engaged in dismantling the tenets of connectionism as an explanatory model for the mind. Many of the tools for that job were already out, in Pylyshyn's (e.g., 1984) and Fodor's (e.g., 1987) own work.

In a seminal paper, Fodor and Pylyshyn (1988) argued that connectionist representations and processes failed to account for some of the key properties of cognitive systems: that they are compositional, productive, and systematic. Crucially, complex representations have constituent structure, which activated nodes in connectionist networks lack. Fodor and Pylyshyn's position on the nature of cognitive architecture has wide consequences for the nature of cognitive representations and processes and, more broadly, for how work on cognitive science ought to progress. Productivity here is key, for if complex representations (thoughts, sentences, perhaps the output of visual processes) do not have constituent structure, are not systematic and, ultimately, if complex expressions are not compositional, then cognitive processes can't be productive. And if cognitive systems aren't productive, how do we manage to say, understand, and think expressions we never said, understood, or thought before? To put it even more dramatically, it seems that the only way to conceive of a mind with an infinite capacity out of its finite resources is to assume that its elementary representations enter into complex structures that are systematic, compositional (and recursive), and thus productive.

It is healthy for any science to have competing paradigms, except that alternatives to symbolic cognitive architecture clearly aren't up to the task. Connectionism cannot account for recursivity, so it appeals to the likes of recurrent networks, which merely mimic recursion. And, as Fodor and Pylyshyn put it, connectionist representations are not compositional: contrary to symbolic expressions, which actually contain their constituent representations, higher nodes that stand for more complex representations do not contain the lower token simplex nodes/representations to which the higher ones dynamically respond. Conversely, a node that stands for a complex representation does not really entail the simplex nodes that are supposed to stand for its constituents. In fact, there is nothing lawful in an association between nodes to the point that a node that stands for P&Q can be associated with P but not with Q. Overall, connectionism cannot give an account of the productivity and systematicity of complex representations: because they are not compositional and do not allow for hierarchical structures and recursion, the only way connectionism can mimic productivity and systematicity is by creating new nodes. But it is not only connectionism that fails to account for the productivity of mental representations: a variety of frameworks (e.g., usage-based language representation, embodied cognition) do too. The main point about Fodor and Pylyshyn's view of the architecture of cognition is that the finite elementary symbols/representations ought to yield for an infinite capacity and the only way known to humankind that this can be achieved is by assuming that cognitive capacities are truly productive (and compositional and systematic), which thus far-circa 2017-only symbolic architectures do.

In his work on cognitive architecture Fodor<sup>15</sup> has emphasized the role of compositionality in complex representations (sentences, thoughts). Compositionality became, in fact, the ensign in the crusade—a "nonnegotiable assumption" in Fodor's take on thought and language. One might suppose that the very idea that the meaning of a sentence/thought should be compositional borders triviality; but it is often the seemingly trivial ideas the ones that make the most noise in cognitive science (take commonsense psychological explanations as a twin example). Compositionality is satisfied, to be clear, when the meaning of a complex expression (sentence/thought) is obtained from the meaning of its constituents (say, morphemes or concepts) and how they are syntactically arranged. As trivial as this might be, opposition to this general principle is the rule rather than the exception. The vast majority of positions in philosophy of language, linguistics, and cognitive psychology, to name the main parties in this dispute, take the meaning of an expression to be rather a function of "semantic features" of the expression's constituents, or to be images, or to be statistical averages (viz., prototypes), or stipulations, or inferential roles, or activation patterns, or to be contextually determined, or something else (the list is vast—and all "or's" are inclusive). How is then Fodor (and colleagues) supposed to take offenders to task? In philosophy of language and mind, Fodor and Ernie Lepore mounted a scathing review of the main positions out there in the market, starting with their Holism: A Shopper's *Guide* (1992). The intricacies of their analyses are way beyond the few words I can write here, but the message is clear: holism is the antithesis of compositionality and thus holism has to be false unless one gives up on the idea that sentences and thoughts are productive and systematic. The crux of the problem goes back to Quine's position on the analytic/synthetic distinction. Since as far as I know nobody has ever came up with the principles for sorting out content-constitutive from contingent properties of a complex representation, the only way to account for lexical-conceptual content while preserving compositionality is to appeal to atomism (of course, contrary to Quine's solution).

Fodor and Lepore's attack on analyticity (of the lexical-content kind) did not stop there: in a series of articles published in the collection *The Compositionality* Papers (2002), they turned their analytical wrath against other offenders. They argued for a position that preserves the "classical" compositionality principle and worked on the details of their approach in typical fashion: showing that a variety of proposals for combining concepts would not work for being committed one way or another to the analytic/synthetic distinction. The solution Fodor and Lepore propose is to assume that lexical concepts are atomic—that is, denotations of token lexical items. Complex representations are obtained only via syntactic/logical form operations introduced by particular types of lexical items. Under their approach a lexical item is complex only in the sense that it specifies, beyond its denotation, a rule for its composition-namely, something akin to an argument structure or a rule for determining the logical form of the expression it partakes. To put it lightly, it's not the content of a token item that is complex, it is its structural/compositional properties—namely, syntax. This view has far-reaching consequences for the nature of semantic/conceptual representations, for the nature of compositionality and, of course, for how language maps onto meaning. With no solutions in sight for the analytic/synthetic distinction, one's choices besides atomism are harsh: either committing to the distinction or abandoning it and adopting some form of holism. Even though these two options lead to a dead end for semantics, a common methodological strategy in the lexical-semantics literature is to shove the problem under the rug and to embark on an empiricist approach to finding the ultimate constituents, the primitives of all lexical concepts.

In several works, notably in *Psychosemantics: The Problem of Meaning in Philosophy of Mind* (1987), A Theory of Content and Other Essays (1990), and

The Elm and the Expert: Mentalese and Its Semantics (1994), Fodor addresses key issues on the nature of content, in particular, on the link between tokens and the properties that concepts express, while mounting a defense of C/RTM for commonsense belief/desire psychology. But it is in *Concepts: Where Cognitive Science Went Wrong* (1998a), where many of these problems are brought to fore in the context of psychological theories. *Concepts* is perhaps Fodor's most developed work on the nature of concepts, and in particular on the metatheory of conceptual representation and development. The book picks up where several other works left off, chiefly *The present status of the innateness controversy*, one of the original chapters of *Representations* (1981). But to get into *Concepts* we need to take a small detour and revisit the early days of lexical semantics.

Fodor's first incursion into the field of lexical semantics (or concepts) was a collaboration with Jerrold Katz, starting when they met in Princeton in the late 1950s and, again, at MIT, in the 1960s. Together they worked on some of the principles of what later became Katz' much more developed semantic theory. In their early work, Katz and Fodor (1963) were strongly committed to a form of lexical-semantic representation that was entirely built on constituent features or "semantic markers." Semantics, for them, was supposed to constitute an autonomous component of linguistic analysis—one that would take the output of structural descriptions provided by syntax and produce a semantic description of token items, based on their constituent features and how they combined. But there was no account of analyticity then, that is, there were no principles governing the selection of semantic markers as constituents of lexical content. And Fodor, soon after, jumped ship.

It is ironic that lexical atomism was borne out of lexical decomposition but that is what happened when Fodor entered into what became known as the "linguistic wars"-though waging a war of his own, one that was not necessarily on the side of the "interpretive semantic" establishment, much less on the side of the opposing "generative semantics." One of the main battles of the "wars" was on the very nature of the division of labor between syntax and semantics: the "generative semantics" movement then assumed that a linguistic description ought to include both syntactic and semantic variables-that putative semantic properties such as causality would constitute part of the grammatical/semantic "deep" constituents that linguistic analyses would yield. The generative-semantics' view, then, was that syntax was not autonomous and that structural analyses of sentences ought to include predicates that were effectively deep-structure representations of surface verbs and their syntactic relations. To put it in other words: the translation of a sentence into its semantic representation required, among its operations, decomposing morphologically simplex verbs into predicate structures containing primitive, morphologically covert predicates (the likes of CAUSE) and their syntactic relations to other sentence constituents. Fodor's (1970) paper, Three reasons for not deriving "kill" from "cause to die" effectively showed that sentences containing the periphrastic cause to die were not synonymous with those containing kill. For instance, we can have "John caused Mary to die on Friday by poisoning her food on Thursday," but not "John killed Mary

on Friday by poisoning her food on Thursday." Unless *cause to die* does not mean *CAUSE TO DIE*, which would be shocking, we should expect the simplex verb and its periphrastic pair to hold their distributional properties—that is, to "behave" the same way—or at a minimum to yield the same semantic representation. But they didn't, so, Fodor concluded, "kill" couldn't possibly mean *CAUSE TO DIE*. It was the end of Fodor's fleeting commitment to semantic decomposition and the beginning of a life-long crusade against it.

The papers that followed, with Merrill Garrett and Janet Fodor,<sup>16</sup> among others, included empirical—namely, psycholinguistic, experimental—investigations of the *kill/cause-to-die* asymmetry and related cases, showing that semantic decomposition does not seem to be at play when we understand sentences. If what we do when we understand sentences is indeed to recover their semantic/ conceptual representations (what else?), we should expect processing complexity effects to arise when simplex verbs by hypothesis turn into complex structures at the semantic or conceptual level. Recall that C/RTM is in effect and more complex computations ought to yield something like greater processing time or some other complexity effect compared to simplex ones. But complexity effects were not obtained in the majority of experiments investigating the semantic complexity of verbs, in experiments that have employed a variety of methods and sentence types.<sup>17</sup>

It is never the case that theoretical advances—or choice between alternatives are solely determined by empirical data. Arguments do carry the heaviest load. In the case of lexical concepts, linguistic and philosophical arguments against decomposition allied to the virtual lack of experimental support for decomposition could be taken as the triumph of the alternative—atomism. Fodor takes up the task of developing atomism more prominently in both A Theory of Content and in Concepts. In this later work, in particular, he looks deep into current (then and now) theories of concepts taking the "nonnegotiable assumption" of compositionality to be the yardstick for measuring the goodness of a concept theory, on the assumption that concepts are the elements of thoughts and that thoughts are compositional. In Fodor's analysis, all decompositional views get similar diagnoses. Concepts can't be *definitions* a la Katz and Fodor or a la Ray Jackendoff and others; definitions can be compositional but, remember, having definitions entail a commitment to the infamous analytic/synthetic distinction, which does not exist at press time. Besides definitions, Fodor's analysis centers on the prototype theory and its kith and kin: concepts can't be prototypes because prototypes do not compose when they enter into complex expressions-that is, they do not contribute their content (their prototypes) to complex concepts, which, by hypothesis have their own prototypes. Think about the PET FISH problem: PET FISH should have its own prototype, which does not have among its constituents the prototypes of PET and FISH. And finally, Fodor shows that if compositionality is to be taken seriously, concepts can't be *theories* either; obviously, theories do not compose and they are at the extreme end of the holism continuum if such a continuum exists. Strictly speaking, holism can't be true because, among a constellation of problems, if our concepts were dependent on all our beliefs, at

a minimum this would violate the publicity of concepts and no two people would ever be talking about the same thing. Moreover, nobody would ever be able to entertain the same thought twice, for the constituents of thoughts would be constantly and forever changing. Neither the publicity nor the stability arguments, of course, deter the proliferation of holistic theories as the current popularity of, say, "embodied" cognition can attest.

Then, if the arguments against holism are right, and if we hold on to the compositionality yardstick, we are left with atomism yet again. It is the only view of conceptual representation that is both compositional and not committed to an analytic/synthetic distinction; the only view of conceptual representation that is compatible with C/RTM. The story seems coherent and well knit, but I am not showing all its knots. The general point is, as Fodor wrote in *Representations*,

If we *are* going to have a cognitive science, we are going to have to learn to learn from our mistakes. When you keep putting questions to Nature and Nature keeps saying "no", it is not unreasonable to suppose that somewhere among the things you believe there is something that isn't true. (p. 316)

The question of decomposition is one for which Nature keeps saying "no." The case against conceptual decomposition—or, conversely, the case *for* atomism— is one in which arguments and much of the experimental evidence point in the same direction. But the last I checked, most concept theories in psychology and lexical-semantic theories in linguistics haven't addressed the key issues that Fodor raised in *Concepts* and in many of the papers that appeared in his *In Critical Condition: Polemical Essays on Cognitive Science and the Philosophy of Mind* (1998b): instead most theories opted for vexingly ignoring arguments against holism, for the impossibility of an analytic/synthetic distinction, and for the central architectural postulate of compositionality. There are sociological arguments for this neglect, but I won't descend to that.

Much of Fodor's subsequent work, including Hume Variations (2003), and LOT2: The Language of Thought Revisited (2008), was dedicated to advancing the cause of C/RTM and making the case for atomism. I say "advancing" but, true to his work, theoretical reflection often involves long and healthy therapeutic sessions (often in group, often in the Insolvent, with Granny, or Aunty, or Snark, or Mr. James, or simply beloved Greycat). The challenges are great, but not insurmountable. For instance, assume that atomism is indeed the only theory compatible with C/RTM and that what C/RTM postulates is that higher cognitive states are essentially relations to propositional attitudes. If concepts are atoms and if atoms are elements of mentally represented propositionsthus, elements of thoughts and their causal relations-how can holism be avoided? In other words, if it is postulated that higher cognitive mechanisms are predicated on the causal relations between beliefs and desires expressed as propositions, on what basis do conceptual/propositional relations obtain? As an admittedly simplified example, consider again the case of *kill/cause to die*. How can the inference x kill  $y \rightarrow y$  die be obtained unless kill is something like

cause to die? Causally determined inferential relations are what functionalism takes to be central to cognitive processes, but the conditions under which inferences are to be obtained appear to be incompatible with atomism, and are surely in conflict with rule-governed, Turing-like computations. Early on, there was an appeal for meaning postulates—a la Carnap—to take care of inferences that appear to be content constitutive. But, in Concepts, Fodor all but abandoned that solution on grounds that meaning postulates that are simply inferences holding between lexical concepts without being *necessary* (viz., encoding empirical knowledge) are, to put it mildly, too weak an alternative.<sup>18</sup> Besides the problems that one faces trying to put together the idea that concepts are atomic with the idea that psychology is intentional and computational, there are problems in the architecture front. As Fodor argues in *The* Mind Doesn't Work That Way (2000), C/RTM (or just CTM) is in trouble for it does not seem to work with abductive inferences, which constitute much of the workings of higher cognitive processes. This is a problem for the architecture of cognition *tout court*—"higher" cognition that is—but not so much for processes that are modular. Something's got to give.

Fodor's perennial existential crisis is the existential crisis of cognitive science it's ours to own. His latest book, with Pylyshyn, *Minds without Meanings: An Essay on the Content of Concepts* (2015), tackles the nature of the connection between the referents—the things out there in the world—and their symbolic representations. Fodor and Pylyshyn take primitive visual attentional mechanisms, the kinds that lock into properties of the world, to establish the causal links between distal stimuli in the "perceptual circle" and their atomic mental representations. Pylyshyn<sup>19</sup> has demonstrated that we attend to and track multiple objects simultaneously and that the connections that are established between the token referents—the things tracked—and their representations are initially "preconceptual." That is, the link serves simply as an individuating mechanism, a form of deixis, as if the visual-attentional system could put its "fingers" on the things tracked or point at them.

Now, let's see what's "inside" the system that affords those links: to begin with, nothing like a "meaning" or an intension (with "s"). In fact, they say it is a "mistake," one that has plagued semantics for about a century (again: *they* say), to identify meaning with intension, following Frege.<sup>20</sup> Here is how they frame the problem: Assume expressions or concepts *JT* (say, Justin Trudeau) and *CPM* (Canadian Prime Minister) both refer to *that* individual *R*. One would imagine that *JT* and *CPM* both hold the same intensional content such that the extension *R* is determined by that content. But as Frege (1892) had shown, the system breaks down in propositional attitude expressions: that the supposed coextension of *JT* and *CPM* does not hold, for an individual can at the same time believe that *JT* refers to *R* while not believing that *CPM* refers to *R*. Fodor and Pylyshyn assume that there is an alternative to the Fregean appeal to intension: since nobody knows what intension is, let alone what a *naturalistic* account of meaning/intension amounts to, it has to go. The alternative is that RTM takes concepts to be "individuated by their extensions *together with their vehicles*" (p. 74).

In other, very rough words, the concept/symbol does not actually *contain* any intensional property, for conceptual individuation is simply a link with its referent. "Meaning is a myth," they proclaim. I'm confident they are not interested in eliminating semantics as a career option, but by claiming that all there is is reference, they are also saying that a lot of the semantic vocabulary—synonymy, paraphrase, translation, and so on—is on its way out.

Turning to the nature of referential links, a key issue is what happens to concepts that are not and never been within the perceptual circle. Those are cases in which Fodor and Pylyshyn take to be the result of long chains of referential connections, cases in which actual referents somehow *were* within the perceptual circle of somebody some time ago. Thinking about Moses, in that sense, implies having a symbol that stands for Moses assuming Moses had been somehow referred to directly sometime, somewhere. Even if we let that pass, for proper names have their own peculiarities, reference to things and events past follow similar chains. Forget "brute force" here: this is more like the case of Plato's "earlier souls" which first triggered the concepts that we now refer by inheritance.

We can't fully evaluate these proposals just yet, not here. While reference within the perceptual circle is well anchored in perceptual and attentional (hence, naturalistic) links, much needs to be said about the representations beyond the "circle," about many concept types and, yet again, about the purported relations between concepts that give rise to categories and other types of inferential processes bearing on the content of propositions. (Quick question: If they don't run on intensions, what do they run on?) But if Fodor and Pylyshyn are at a minimum half-right, the cognitive science of concepts will be required to do some work on its foundations, much like their missing epigraph would have recommended,

> If you slip . . . Pick yourself up Dust yourself off And start all over again (Jerome Kern & Dorothy Fields)

**F** or long our belief boxes have been holding a symbolic expression meaning that Fodor has been the most prominent figure in some of the most important battles leading to cognitive science's current stage, to its autonomy from behaviorism and physicalism, and for its focus on the nature of mental representations and processes. He has set the agenda for some of the most important debates shaping the core of the field—from the nature of cognitive architecture to the nature of concepts. One certainly can't tell what would have been of cognitive science, of its second revolution, without some of the metonymic names fighting its most important battles against behaviorism (then and now) and against the reductionism that physicalism (then and now) promotes. And one doesn't know in particular what would have been of all this without Fodor. But there is no doubt about what happened to the field when he came into play. Sometimes battles are fought alone, sometimes under quixotic delusions, as the knight in the well-known story put:

Fortune is arranging matters for us better than we could have shaped our desires ourselves, for look there, friend Sancho Panza, where thirty or more monstrous giants present themselves, all of whom I mean to engage in battle and slay. . . (M. de Cervantes, *Don Quixote*).

It just so happens that sometimes windmills *are* indeed giants worth slaying. In Fodor's case, there were giants, the targets of his unique analytic mind, some of whom still linger despite being inflicted with mortal arguments. Nobody knows what will be of Fodor's work 300 years from now (assuming exceptions are made, I shall update this guide). Descartes' contributions to philosophy are still at the forefront of the debates on how the mind works. Hume's work was, on his own account, initially "overlooked and neglected,"<sup>21</sup> but look at him now. We do know that Fodor's impact has been immediate upon entering the cognitive science scene—and that he has been engaging and slaying giants ever since. In the process, anarchic as it has been, the cognitive revolution achieved many of its goals. Old Granny does not visit anymore, though her psychographic messages keep recurring in connectionist writings. History—fairly or unfairly—will hold Fodor as a metonym for the kind of cognitive science that was, is, or ought to be.

## AUTHOR'S NOTE

I plead guilty to false advertisement, for I do not—and cannot—provide anything near a *complete* guide to all the many attractions. What is provided here is a very rough map to some of the issues that have occupied Fodor's mind and have helped set the agenda for cognitive science. I also limit the scope of the discussion to the topics that have occupied the minds of the editors and contributors to this volume, undeniably under Jerry Fodor's spell. Even the title of this introduction is, of course, inspired in the title of one of Fodor's papers ("Fodor's Guide to Mental Representations," 1985; millennials are supposed to Google "Fodor's guide" to get the joke). We—Lila R. Gleitman and I—are certainly most grateful to Jerry for all. I am also indebted to Caitlyn Antal, Tom Bever, Noam Chomsky, Lila R. Gleitman, and Ernie Lepore for comments on earlier versions of this chapter, and to the National Sciences and Engineering Research Council of Canada (NSERC), for support.

### NOTES

- 1. I occasionally use "we" to refer to both editors of this volume or as a generic pronoun.
- 2. See Miller, Galanter, & Pribram (1960) and their interest in exploring "cybernetic ideas" in psychology—specially "plans" as cognitive programs. These "cybernetic

ideas" were well under development in the 1950s (see, e.g., Newell, Shaw, & Simon, 1958; and the papers in Feigenbaum & Feldman, 1963).

- 3. A product of this visit was Fodor (1965), an analysis of behaviorists' account of meaning as "mediating" responses.
- 4. In Rieber (1980, p. 80).
- 5. See chapters by Bever and Garrett in this volume.
- 6. See, in this volume, chapters by J. D. Fodor, Nickels, & Schott and by Ferreira & Nye.
- 7. We could just as well take the symbolic level to be part of biology. Here I yield to convention and treat them as separate levels of analysis.
- 8. See, for instance, Gallistel's chapter in this volume.
- 9. For ease of exposition, I will collapse two theses, RTM and CTM. You can be committed to the idea that there are representations of some sort without being committed to the idea that processes over those representations are computational, Turing-like. If you are committed to the latter, you have to be committed to the former, and that commitment in turn restricts the nature of representations (viz., to those that are computable). For the most part, Fodor is committed to both, but see his *The Mind Doesn't Work That Way: the Scope and Limits of Computational Psychology* (2000), where he discusses varieties of CTM, and why he assumes that CTM, strictly speaking, only holds for modular processes typical of input systems—not holistic ones, typical of central-system processes. I return to this later in the discussion on modularity. See also, de Almeida & Lepore (this volume).
- 10. If a mosquito bites you, most likely the cause of your desire to scratch the itch you got—and ultimately whether or not you actually scratch yourself—is not computationally derived, not in any sense that, say, the conclusion in a *modus ponens* is.
- 11. See, in this volume, the chapter by Lobina & Garcia-Albea, on the relation between LoT and the faculty of language.
- 12. See chapter by Piattelli-Palmarini in this volume. See also Fodor & Piattelli-Palmarini's *What Darwin Got Wrong* (2010), where Darwin's natural selection theory is taken to be analogous to behaviorism's learning theory, presupposing nothing in terms of the organism's internal states in the process driving evolution.
- 13. See chapters in this volume by Chomsky, Garrett, Ferreira & Nye, de Almeida & Lepore, Pylyshyn, and Potter.
- 14. See, in particular, Fodor (1983, 2000) and Chomsky's chapter in this volume. For an early treatment of Chomsky's notion of modularity, see Chomsky (1980).
- 15. Besides Fodor & Pylyshyn (1988), see also Fodor & McLaughlin (1990).
- 16. See J. D. Fodor, Fodor, & Garrett (1975) and Fodor, Garrett, Walker, & Parkes (1980).
- 17. I say "majority" because there have been a few experiments claiming to support verb-semantic decomposition, all of which face some harsh problems. A recent review of these appears in de Almeida and Manouilidou (2015).
- 18. The idea that there are non-content-constitutive meaning postulates is not necessarily a weak, unconstrained alternative; it might be simply the best one can get out of rule-like processes in an otherwise holistic environment, thus at least preserving a weak version of CTM without being committed to "inferential role semantics." But this cannot be worked on here (see de Almeida, 1999, for an early attempt).
- 19. See Pylyshyn's chapter in this volume.

- 20. The reader might want to brush up on the so-called Frege cases (viz., "the morning star" and "the evening star" as both referring to Venus; and the problem posed by the use of these expressions in propositional attitude statements) and, on the way back, to look at Putnam's case (the Twin Earth argument). Both types of cases have been subject to Fodor's scrutiny (see, e.g., Fodor, 1987, 1994, and 2008). It should be noted that neither Frege nor Putnam takes meaning to be "in the head." Fodor's reading is that at least in Frege's case expressions or concepts are token mental representations—that, e.g., *THE MORNING STAR* is a concept, in fact a different one from *THE EVENING STAR* even though both refer to Venus.
- 21. This refers to the reception of his *Enquiry Concerning Human Understanding*. See Hume's (1777/2009) *My Own Life*.

## REFERENCES

Chomsky, N. (1959). A review of BF Skinner's Verbal Behavior. Language, 35(1), 26–58.

- Chomsky, N. (1980). *Rules and representations*. New York, NY: Columbia University Press.
- de Almeida, R. G. (1999). What do category-specific semantic deficits tell us about the representation of lexical concepts. *Brain and Language*, *68*, 241–248.
- de Almeida, R. G., & Manouilidou, C. (2015). The study of verbs in cognitive science. In R. G. de Almeida & C. Manouilidou (Eds.), *Cognitive science perspectives on verb representation and processing* (pp. 3–39). New York, NY: Springer.
- Feigenbaum, E. A., & Feldman, J. (1963). *Computers and thought*. New York, NY: McGraw-Hill.
- Feyerabend, P. (1975). *Against method: Outline of an anarchistic theory of knowledge*. New York, NY: Verso.
- Fodor, J. A. (1965). Could meaning be an rm? *Journal of Verbal Learning and Verbal Behavior*, 4(2), 73–81.
- Fodor, J. A. (1968). *Psychological explanation: An introduction to the philosophy of psychology*. New York, NY: Random House.
- Fodor, J. A. (1970). Three reasons for not deriving "kill" from "cause to die." *Linguistic Inquiry*, *1*(4), 429–438.
- Fodor, J. A. (1975). The language of thought. New York, NY: Thomas Y. Crowell.
- Fodor, J. A. (1981). *Representations: Philosophical essays on the foundations of cognitive Science*. Cambridge, MA: Bradford Books/MIT Press.
- Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: Bradford Books/MIT Press.
- Fodor, J. A. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind.* Cambridge, MA: Bradford Books/MIT Press.
- Fodor, J. A. (1990). A theory of content and other essays. Cambridge, MA: MIT Press.
- Fodor, J. A. (1994). *The elm and the expert: Mentalese and its semantics*. Cambridge, MA: MIT Press.
- Fodor, J. A. (1998a). *Concepts: Where cognitive science went wrong*. New York and Oxford, England: Oxford University Press.
- Fodor, J. A. (1998b). In critical condition: Polemical essays on cognitive science and the philosophy of mind. Cambridge, MA: MIT Press.

- Fodor, J. A. (2000). The mind doesn't work that way: The scope and limits of computational psychology. Cambridge, MA: Bradford Books/MIT Press.
- Fodor, J. A. (2003). Hume variations. Oxford: Clarendon Press/Oxford University Press.
- Fodor, J. A. (2008). *LOT 2: The language of thought revisited*. Oxford, England and New York, NY: Oxford University Press.
- Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). The psychology of language: An introduction to psycholinguistics and generative grammar. New York, NY: McGraw-Hill.
- Fodor, J. A., Garrett, M. F., Walker, E.C. T., & Parkes, C. H. (1980). Against definitions. *Cognition*, 8(3), 263–367.
- Fodor, J. A., Lepore, E. (1992). Holism: A shopper's guide. New York: Wiley-Blackwell.
- Fodor, J. A., & Lepore, E. (2002). *The compositionality papers*. Oxford, England: Oxford University Press.
- Fodor, J. A., & McLaughlin, B. (1990). Connectionism and the problem of systematicity: Why Smolensky's solution doesn't work. *Cognition*, *35*(2), 183–204.
- Fodor, J., & Piattelli-Palmarini, M. (2010). *What Darwin got wrong*. New York, NY: Farrar, Straus and Giroux.
- Fodor, J. A., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28(1–2), 3–71.
- Fodor, J. A., & Pylyshyn, Z. W. (2015). *Minds without meanings: An essay on the content of concepts*. Cambridge, MA: MIT Press.
- Fodor, J. D., Fodor, J. A., & Garrett, M. F. (1975). The psychological unreality of semantic representations. *Linguistic Inquiry*, 6, 515–531.
- Frege, G. (1892). On sense and reference. In P. Geach & M. Black (Eds.), *Translations from the philosophical writings of Gottlob Frege*, 2nd ed. (pp. 56–78). Oxford, England: Basil Blackwell.
- Hume, D. (1777/2009). My own life. In D. F. Norton & J. Taylor (Eds.)(2009). *The Cambridge companion to Hume*, 2nd ed. (pp. 522–529). Cambridge, England: Cambridge University Press.
- Katz, J. J., & Fodor, J. A. (1963). The structure of a semantic theory. *Language*, 39(2), 170–210.
- Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York, NY: Holt, Rinehard & Winston.
- Newell, A., Shaw, J. C., & Simon, H. A. (1958). Elements of a theory of human problem solving. *Psychological Review*, 65(3), 151–166.
- Pylyshyn, Z. W. (1984). Computation and cognition: Toward a foundation for cognitive science. Cambridge, MA: MIT Press.
- Rieber, R. W. (Ed.) (1980). *Dialogues on the psychology of language and thought*. New York, NY: Plenum Press.