LANGUAGE OF THOUGHT AND MENTAL LOGIC

Majid Amini and Niloufar Jafari

Jerry Fodor's name has become almost synonymous with several theoretical theses in the interdisciplinary field of psychophilosophy, one of which is the *language of thought* published in a book of the same title in 1975 that later he modestly described as a 'merely journalistic' enterprise: "an exposition of . . . an emerging, interdisciplinary consensus about how the mind works; the theory that was then just beginning to be called 'cognitive science'" (2008, 3-4). The hypothesis, to put it succinctly, states that there is an *innately endowed* code in which mental processes are carried out. Various arguments have been presented in support of this hypothesis, and what is interesting is that almost all of them bear an *empirical* facet, except possibly for one.[1]

Apparently, there is one *a priori* argument that seems to have been advanced by Fodor essentially in point of logic. Fodor's reasoning would go something like the following: being able to represent involves possessing a means in which representations can be expressed; since the best psychological theories are representational, one is therefore bound to accept the existence of an internal code (Fodor 1976).[2] Appearances aside, even this argument is not purely *a priori* in nature due to its empirical appeal to the premise that the best psychological theories are representational. Indeed, for this very reason, on a number of occasions, Fodor has attempted to evince that various empirical theories need explicit representations.

In support of his hypothesis, Fodor also adduces other considerations, though still empirical in alignment, like the claim that the best theory satisfying plausible *desiderata* for propositional attitudes involves an internal language (Fodor 1981). All in all, the reasons advanced for the language of thought were mostly empirical in one way or another. Clearly, the situation could not have been otherwise as a representational theory of mind is an empirical thesis after all.

The hypothesis, however, has not gone unscathed since its formulation, and detractors

---

[1]Notwithstanding, however, some subsequent attempts to establish the hypothesis on *a priori* or conceptual grounds like Davies 1991, Lycan 1993, and Rey 1995. However, Fodor is adamant that the constituent components of the language of thought hypothesis are "at best *empirical* theses" (2008, 93).

[2]To avoid misrepresentation, a caveat is in order here. It is often stated, for example, by Elliot Sober among others, that "cognitive processes are to be thought of in terms of inferential transformation of mental representations. . . . Those psychological models which now seem to be promising presuppose that perceptual judgment, memory, and language acquisition, to name some examples, take place within representational system ('the language of thought') in which information is enclosed and manipulated" (1981, 97). In other words, representation, as Robert Cummins notes, is "identified with what is really only one kind of representation: quasi-linguistic representation of the sort featured in Fodor's book *The Language of Thought*" (1989, 18). However, strictly speaking, this is not correct since cognitive states might involve representations of some sort without involving a language of thought. It should be said that occasionally Fodor's own way of presentation is to some extent responsible for encouraging such identifications, where, for example, he says that "I'm much inclined to bet on a species called the Language of Thought Hypothesis . . . or, alternatively, the Representational Theory of Mind" (1988, xii).

Amini, Majid, and Niloufar Jafari. "Language of Thought and Mental Logic." *Shibboleths: a Journal of Comparative Theory and Criticism* 5 (2017-2018): 17-28.

have been quick in exposing its vulnerabilities. The increasingly stronger criticisms have showed, if nothing else, that the language of thought hypothesis is far less obvious than originally thought. Yet, the barrage of attacks would still be to the advantage of the hypothesis, should it turn out to be true. The purpose of this discussion is to consider the major criticisms of the hypothesis against the background of another *nativist* doctrine, *viz.* theory of *mental logic*, along with its empirical implications in an attempt to provide another source of evidence for the hypothesis, and possibly to shed some light on the structure of the language of thought. Clearly, the intention is to show that the criticisms play more a corrective role rather than being detrimental to the hypothesis.

At this juncture, therefore, it would not be amiss to state a brief description of mental logic theory.[3] The theory maintains that deductive reasoning consists of innate operations on internal representations in accordance with logical rules implemented in procedures activated by the forms of mental representation. The underlying thought here is that the set of valid inferences is infinite and the subset of those that any individual will formulate or encounter in the reasoning of others is too large and too random to permit its being memorized. Humans, thus, appear to have access to a set of rules that can be combined in various ways to yield an infinite set of inferences. This means that the foundations of the logic(s) at which logicians aim, *viz.* the logical precepts and ideals, must be psychologically real in the sense of being instantiated in some form in the mind. It is thereby contended that this theory has the ability to satisfy the twofold requirement of offering the best explanation for everyday deductive reasoning as well as delivering some insights into human cognitive architecture.

As intimated earlier, the language of thought hypothesis is said to rely heavily on the doctrine of *representationalism*, but it is contended that the alleged evidence for it does not exist and the doctrine lacks empirical justification. This is one of the major criticisms that challenges the main premise common to all the arguments for the language of thought hypothesis. That is, it is not true that the best computational theories require explicit representations. The objectors point out that a large part of the original motivation for a language of thought was to make sense of how systems of rules such as the ones postulated by linguistics could be true of organisms like us. The proposed suggestion was that rules are explicitly represented and as such they are represented in the language of thought. But, the objectors claim, in fact no evidence supporting the view that the rules are explicitly represented has been really offered. The criticism is essentially saying that even if one knows the input-output relations of a cognitive function, one still lacks a sure understanding of what happens in the intervening black box where representations of rules should be needed. Robert Matthews, for example, concludes thus:

> Yet in each case the evidence adduced falls short of establishing the claim in question, and always for the same reason: the evidence adduced bears only on the input/output behavior of speakers or learners; it provides no support for hypotheses to the effect that such behavior is the product of a representation-using system. It could as well be the product of an input/output equivalent non-representation-using system. (1984, 1083)

It has even been claimed that no such evidence could be in principle provided as a matter of logic. This stronger rejectionist response is predicated on the consideration that it is a

[3]To name just a few luminaries that pioneered and produced works in the earlier stages of mental logic as a distinctively nativist doctrine in contrast with non-nativist positions like mental models hypothesis, one may mention Mary Henle (1962), Daniel Osherson (1975 and 1977), Martin Braine (1978 and 1990), and Lance Rips (1983 and 1994).

theorem of recursion theory that the same computable function has an infinite class of different indices, or names, associated to it; in particular, if in this class there is a program containing an explicit representation of a rule, there will also be another program computing the same function but not containing that rule explicitly represented. The point is equally applicable to grammars, for given a device which contains an explicit representation of a grammar, it is always possible to find another device that computes the same function but does not contain that grammar. Rules, therefore, need not be represented and by the same extension the need for a structured internal code becomes unmotivated.

It should, nevertheless, be obvious that the strong rejectionist version of the attack on *representationalism* is very much a statement of *conventionalism*. This, however, should have already rung the alarms, since the argument can easily be applied to the whole of science and the entire scientific fabric. Apart from such concerns about the integrity of science, the argument in both versions, *viz.* evidential poverty and logical impossibility, still misses the target. For, even if one grants that there is yet to be a clue to the internal representation of the rules/procedures of a cognitive operation, *already the form of their input/output data may suffice to make the case for representationalism*. It is here that mental logic as a theory of deductive reasoning can offer some headway towards vindicating representationalism. Even if one concedes that there is no need for one to hold the explicit representation of the rules of mental logic, one yet could consistently maintain that the data on which a reasoning program operate require a language of thought. The possibility becomes plausible once one looks at some basic facts about deductive reasoning.

The first thing to note is that the input to reasoning *cannot* be in a format somehow dependent on a single natural language. The reasoning here is twofold: *homophonic* and *heterophonic*. In the homophonic case, there are far too many ways of expressing the same proposition in a natural language, and as such for the same inference one would need as many different inference patterns as there are ways of expressing the component propositions in the native language. But this seems unjustifiably extravagant. With regard to the heterophonic case, reasoning in, for example, Persian is the same as reasoning in, say, Portuguese which precludes the dependence of reasoning on a single natural language format. These considerations are equally true of linguists' *logical form*, since so long as it is natural language-oriented or dependent, it will not have the right level of abstraction. As Robert May observes, although to a large extent properties of logical form seem to be invariant from language to language, "one could imagine a range of differences as a function of independently varying properties of the S[Surface]-structure input to the rules deriving LF [Logical Form], for instance" (1985, 30; 1987, 332).

The second point to consider is that direct reasoning routines are at least early available and hence, presumably, species-specific (Braine and Rumain 1983; Braine, *et al.*, 1984; Braine 1994). For example, disjunctive and conditional phrases appear fairly early in development, typically during the third year, over several natural languages investigated; while, conjunction and negation appear even earlier. What is important to note is that when they appear, they are almost always used in a semantically appropriate way, indicating that children have an early grasp of their meaning.

Thirdly, the way mental logic procedures modify the data requires that they be formulated following a grammar that respects constituency relations. For example, if a conclusion is in conditional form, the proposed program of mental logicians like Braine and company will add its antecedent to the premise set, *but not* any arbitrary subformula or one of its incomplete parts. Also the antecedent and consequent of the conditional sentence may themselves be complex sentences, but generally the program sees their complexity only when

they have been broken up by some procedure acting on the main connective of their logical form. Consequently, the way data are stored in working memory and accessed by the routines of the program must maintain constituency and dependency relations among connectives. In fact, as Braine observes,

> some sort of innate format for representing knowledge would seem to be a precondition for memory, at least for declarative memory . . . if there is no format for recording something, there is no way to record it. Now, to be adequate to the task of recording information in memory, the format would need some logical structure. For example, it would need some kind of predicate-argument structure in order to distinguish properties from the entities remembered to have the properties, and relations from the objects noticed to be related. (1994, 243-244)

For the same reason, it would also need to be able to express connectives and their hierarchical dependence in a sentence.

And, for the fourth fact, it should be said that the input to a reasoning program is not the first level of semantic analyses of a linguistic signal, but a composed representation maintaining the meaning of the message in context. This fact would also be indicative of the intervening of an internal code in which reasoning processes take place.

Now, collectively all these considerations point towards the requirement that the input data for a reasoning program be written in a language different from natural language. In other words, there needs to be a language with a regimented grammar defining conditions for well-formedness as well as being sufficiently rich to express all the distinctions of, for instance, categories or constituents, to which the reasoning program would be sensitive when drawing inferences. That is, there needs to be a language of thought.

Now, if it is independently shown that mental logic as a theory is both *descriptively adequate* and *psychological real*, the existence of an internal code and some features of its structure can then be inferred from the *data structure* alone. This would obviously be contrary to the above two arguments against representationalism and indicates how they have managed to miss their intended target.

However, there is still no cause for celebration for the advocates of the representational theory of mind, since the conflict has just begun and there are other more caustic criticisms to contend with. One such objection to representationalism argues that although the hypothesis is empirical in orientation, it is at the same time empirically too vague to the extent that its claims to solve the problems it purports to solve are rather illusory and elusive. For example, Matthews, among others, claims that "instances of notably successful cognitive scientific theorizing that have been claimed by Representationalists to provide empirical support for their doctrine" (1989, 104) *either* fail to provide any essential role for the theory *or* the roles on offer "cannot plausibly be construed along the lines" (1989, 104) of the representational theory of mind.

It is remarked that more often than not the arguments for accepting the language of thought hypothesis depend on some or other of its alleged properties, but at the same time there is no evidence that the hypothesis *has* the property in question. Even if one grants that it does, other substantial worries arise about the structure of language of thought for which there is no clear answer. Apparently, closing a hole in one place opens a crater in another part of the hypothesis such that its advantages may be heavily overwhelmed by its shortcomings.

The concern is very genuine especially when one considers the three arguments on which so much 'propaganda' for the representational theory of mind is based, namely, the

supposed parallelism between syntax and semantics, the systematicity of thoughts, and the opacity of propositional attitudes. The arguments have all been independently presented to prove the advantages of a language of thought hypothesis. But, the opponents of representationalism ask, are they simultaneously satisfiable? They have tried to capitalize on the alleged conflicts and inconsistencies to undermine the concept of a language of thought.

Take, for example, the arguments from the parallelism between syntax and semantics and the systematicity of thoughts. It is claimed that the language of thought theory is able to show jointly the parallelism between semantic relations among intentional contents and causal mental processes and why thoughts are systematic. Obviously, the two issues are tied and interrelated, but, the opponents of representationalism point out, there is a significant amount of details that need to be spelled out before accepting such assertions. And it is only in the details that one could see whether the language of thought hypothesis can really offer an explanation for them. But, it is claimed, the details are yet to emerge.

This is, however, a vague criticism since, as Fodor notes, we need a language of thought because Alan Turing's explanation is the only explanation and for that we need a syntactic vehicle for each semantic content. Indeed, programming languages are the best existence proofs to vindicate the propriety of the language of thought: "[c]omputers show us how to connect semantical with causal properties of *symbols*" (1988, 18). Nonetheless, to see the import of uneasiness, one may concentrate further on the notion of parallelism.

The first thing to note about the parallelism thesis is that it can be predicated at *different* levels. At one level, it could be *inter-sentential*, according to which propositions are structured objects resulting from the composition of entities of different types where the parallelism thesis states that *so are* thoughts. This is the way that Kant and Frege saw the issue. That is, the thought that Adam is untidy presumably contains two tokens of different categories, the name 'a' and the predicate 'U', composed according to the predication rule into 'U(a).' The structure of thought mimics the semantic of the proposition, since the proposition that Adam is untidy consists of a property attributed to an object. To satisfy inter-sentential parallelism, therefore, the language of thought must have symbols for different categories, which in their turn have different semantic values.

The question, therefore, is: how can one find out whether the parallelism really holds, if there is no syntax to begin with? One way of eschewing the stalemate is to find *a theory* capable of specifying what categories exist and what laws of compositions are defined etc. Armed with such a theory, one would then be in a position to say that the language of thought hypothesis *explains* inter-sentential parallelism. Only under such circumstances can one ascertain whether it has as consequences properties of thought like systematicity.

However, we are not out of the woods yet. For, specifying conditions of well-formedness is not sufficient to explain inter-sentential parallelism, because it would be pointless to have a syntax of thought which can maintain in phase single propositions and single thoughts, but lose the contact with meaning when propositions are chained. To illustrate the issue, consider the following example which is frequently forwarded as a proof of the explanatory power of the language of thought hypothesis. The proposition that P and Q is said to be semantically related to the proposition that Q: that is, if 'P & Q' is true, then 'Q' is also. It is stated that the language of thought theory would capture such connections by possessing a rule according to which the symbol 'P and Q' is syntactically tied to the symbol 'Q'. Yet, the parallelism is once again preserved by introducing a piece of syntax in the language of thought.

Now, for one thing, although in *logic* such a parallelism can be thoroughly respected, it is an *empirical* issue whether some of the same logical rules have counterparts in the mind;

and, for another, will the additional specification of transformation of symbols in the language of thought to ensure that they *are* truth-preserving suffice for the explanation of parallelism? As regards the latter, apparently not. For, one must also make sure that these features and properties do not clash with the claims of other arguments for a representational theory of mind. In particular, it should cohere with the claim of representationalism being able to account for propositional attitudes, *inter alia*, their opacity. This is the third and one of the most often repeated arguments for the representational theory of mind.

To see whether the foregoing conditions on the language of thought square with the claims of the third argument, the problem of *opacity* of propositional attitudes needs to be addressed. Yet, the opacity of propositional attitudes and other associated issues have turned out to be rather intractable and, indeed, a constant source of inspiration, if not irritation, for theories of propositional attitudes. In fact, even elaborate relational theories like Rudolf Carnap's *intensional isomorphism*, developed and refined, for example, by David Lewis, run into difficulties (Carnap 1988, 56ff.; Lewis 1972, 182ff). For example, allowing substitutions of intensionally isomorphic beliefs could still lead to problems like someone believing that the Greek Tourism Office is closed without believing that the Hellenic Tourism Office is closed, where the two beliefs are not only intensionally equivalent, but also intensionally isomorphic. That is,

> this approach does not help distinguish a pair of belief sentences where the embedded clauses have the same syntactic structure and differ only in that one has a synonymy of the corresponding word in the other . . . since the meanings of the embedded clauses (in this technical sense) will still be the same. (Dowty, *et al.* 1981, 173)

Now, can a theory of language of thought contribute in any way towards some solution to this problem? It seems that by maintaining the relational character of beliefs, yet adding to the relation a place for the syntactic counterpart of a proposition in the language of thought, one can apparently obtain the right degree of opacity necessary to block logical omniscience as well as many other unwanted logical consequences of held beliefs.[4] The belief that P is not the same as its logically equivalent belief that ~~P; the belief that Hesperus is a star is not the same as its semantically equivalent belief that Phosphorus is a star; and, the belief that triangles are equilateral is not the same as the intensionally isomorphic belief that triangles are equiangular.[5] For, in all cases, at least one member of the relation, *viz.* the syntactic parameter, is different.[6]

To give an interim conclusion, it seems that in order to account for failures of substitutivity in propositional attitude verbs, representational theory of mind proposes to

---

[4]Despite a consensus, there are philosophers who take exception to it. For example, Matthews attempts to sketch the "broad outlines of a non-relational measurement-theoretic account of propositional attitudes" (1994, 131). However, notwithstanding his animadversions towards the theory of language of thought, Matthews readily admits that his alternative account is still "compatible with language-of-thought hypotheses that attribute to propositional attitudes the semantic, syntactic, and inferential properties of the representation space into which they are mapped" (1994, 144).

[5]Assuming rigid designation, of course!

[6]It is also possible to extend the treatment to puzzles like the ones by Saul Kripke (1979, 119ff). By exploiting the presence of a syntactic parameter in the analysis of propositional attitudes, Fodor shows how to deal with cases where the belief that London is pretty is not the same as the belief that *Londres* is (1990, 168ff.; 2008, Chapter 3).

introduce a syntactic parameter in the relations specifying the relevant attitudes. However, this is not the end of the matter, since not *any* theory will do if one notes that there is also the opposite problem. To add syntax as an independent parameter for identifying beliefs may give a way to solve the problem of their opacity, but it may open another, perhaps deeper, problem: namely, the problem of making psychological laws predictively too poor. Generally, laws are in conditional forms, issuing certain predictions once the antecedents of the conditional are satisfied, and the antecedents typically contain references to people's mental states. Now, if the syntactic parameter in the intentional relations is the factor responsible to block the unwanted inferences, and the laws are sensitive to it, then one may risk the ability of predicting that two persons fall under the same laws only when they have tokens of the exact same type in their heads. The reason is that as every syntactic difference potentially counts as a difference in beliefs, the proposed theory *in principle* may block *all* the substitutions. Yet, one would like to be able to predict that Adam will leave the building if he believes that it is in flames, and Eve will leave it as well if she believes that it is burning, *regardless* of whether the actual token is of the type 'the building is burning' or 'the building is in flames.' The moral is that the theory *must* allow for a class of substitutions, but not for *all*, and here is where the difficulty lies. Since, once again, one does not know which substitutions the theory licenses unless the laws of derivations governing transformations of symbols in the language of thought have already been specified.

The story of the language of thought so far is like this: it is offered as a solution to the question of systematicity, to explain the parallelism between syntax and semantics, and to deal with the opacity of the attitudes. It was argued that in order to account for systematicity and for inter-sentential parallelism, the language of thought hypothesis must be supported by a description of its categories and of its well-formedness rules. For, without them, it does not seem able to shed any light on those issues. However, in order to account for inter-sentential parallelism, the theory must also specify the laws connecting sentences, *i.e.* it must be supported by a theory of reasoning. But, if the language of thought is going to keep all its promises, it also becomes clear that, even in this case, not every theory will do.[7] That is, it transpires that laws connecting sentences must be specified in order to account for parallelism. Yet, whether the language of thought hypothesis can also simultaneously solve the problem of the opacity of attitudes depends on how such laws turn out to be. In other words, if they are too liberal in the inferences they allow, then the purported advantages of the language of thought disappear, and if they are too strict, then the language of thought will turn out to be empirically inadequate to account for the inferences and substitutions that must be allowed for psychological categorization. In short, the naked existential statement that there is a language of thought is a blank cheque which can be cashed only when one actually reads the figures written on it. This means that to put a *lower bound* condition is not sufficient, and the explanatory power of the hypothesis is very much in the thrall of the details of its development.

Yet, it should be said that the history of the development and emergence of the hypothesis has not been that conducive towards assuaging the suspicions of the critics either. In his first substantial work on the language of thought, Fodor undertook to frame at least the main lines of a theory for it by trying to gather information about its structure from empirically motivated independent theories of cognition (Fodor 1976). However, what subsequently transpired was that the defense of the hypothesis shifted towards more indirect strategies,

[7]Obviously, this is not a criticism but just to highlight the additional empirical aspect of the situation.

and, for better or worse, such strategies are unlikely to throw a great deal of light on the structure of language of thought, thus making the sceptics more bullish and suspicious (Fodor 1988). The indirect evidential approach seems still to be in force in the last direct contribution of Fodor on the language of thought hypothesis (Fodor 2008).

The opponents have been thereby quick in latching onto the *empirical underdetermination* of the language of thought theory and have attempted to exploit the situation to its full mileage. The problems posed by the critics could be summarized in the form of three main questions. The first question concerns the issue of why inferential practices of subjects are not closed under rationally acceptable rules of deductive inference if the hypothesis is correct. That is, although one may talk about performance failures, memory limitations, interaction effects etc., it is incumbent on one to construct an account that actually predicts and explains observed inferential practices at some point. Secondly, how can one overcome the problem of specifying the syntax of the language of thought? The second question, in turn, leads to the third issue that without the specification of a syntax for the language of thought, there is no clear and concrete proposal for the finite encoding of what advocates of the hypothesis presume to be an infinite number of beliefs. To give an actual quote, Matthews, for example, complains that,

> [v]ery little is offered by way of a characterization of the semantic relations that the causal relations are said to respect; nothing is offered by way of a characterization of the syntax in virtue of which the symbol tokens postulated by the RTM [Representational Theory of Mind] are said to interact causally. (1991, 142)

Notwithstanding the foes of the language of thought hypothesis, even its friends have not shied away from expressing the same sentiment. Braine, for instance, remarks that,

> [a]ctually, Fodor has very little to say about the syntax of the language of thought, although it is its syntax that defines it as a system of representation, and in which its logical properties are embodied. (1994, 244)

In other words, there has been an intriguing disconnect and disengagement between these two prominent nativist doctrines of mind. Thus, in pursuit of a meaningfully direct and collaborative interaction between these two nativist hypotheses, one may draw on the resources of mental logic to see whether it is able to transform the language of thought hypothesis in much more than a naked existential way by making an attempt to address the above misgivings and shortcomings.

To begin with the syntax problem and the connected issue of finite encoding, it is obvious that lack of the specification of a syntax for the language of thought is cause for concern. Yet, it would be unfair to ask from the hypothesis anything more than what is asked from other sciences. For, a large part of the syntax of natural language and of its semantics is not yet determined, but this is no reason to refute the current theories. What, generally, linguists attempt to do is deploy a tactic of *divide and rule* by which they try to cover syntax and semantics for increasingly more comprehensive fragments of the language under investigation. In the same way, mental logic allows one to apply the same strategy to the language of thought.

Mental logicians have been able to provide analyses of fragments of reasoning such as propositional as well as predicate and quantificational reasoning such that they can be drawn on as pieces of the syntax of the language of thought and offer information about some of the categories of basic symbols and the rules of composition that a mind must possess in order to perform even elementary deductive problems (Braine and O'Brien 1998; Rips 1994). They provide fragments of the finite encoding of an infinite generative capacity that critics like

Matthews demand. More importantly, they provide methodologies and strategies to extend the investigation to other domains.[8] What is, however, significant to observe is that lack of a full determination of the syntax of language of thought is not an insurmountable obstacle, since the existing data already support conclusions about the minimal generative capacity that the system must possess in order to feed the reasoning schemata with complex sentences. Therefore, mental logic seems capable of taking care of the last two major problems that are often hurled at the language of thought hypothesis.

Now, can mental logic also take care of the first major problem posed for the language of thought theory, namely, to explain why inferential practices of subjects are not closed under rationally acceptable rules of deductive inference, if the hypothesis happens to be true as alleged by its advocates? Before making any attempt at an answer, if one, however, looks at the question more carefully, there are, in fact, three issues that have been rolled into one query. The threads would therefore need to be separated out first.

One issue directly concerns subjects' logical competence and indirectly the purported parallelism between syntax and semantics. If the reasoning system mirrors semantic connections among intentional objects, then a language of thought supporter should be able to explain why, quite aside from performance limitations, subjects are prone to comply with a valid inference in certain cases and not in others. Apparently, the inferential mechanism and the semantic connections it purportedly mirrors somewhere, must take separate routes, and thus the language of thought must indicate exactly where and how this separation takes place. Therefore, what is demanded of the hypothesis is to spell out the *closure conditions* for the supposed inferential mechanism. The second thread in the question concerns the *use* of a *competence/performance* distinction. The point is simply that if supporters of language of thought see fit to appeal to performance failures in attempting to account for some apparent violation of rational behavior, it well behooves them to show that their usage of the distinction does *not* beg the question. In other words, any appeal to such a distinction is made *in principle* and not just to avoid embarrassing putative counterexamples. Lastly, the third thread of the question consists of a demand for a theory with adequate predictive power, that is, to present an account that reliably predicts observed inferential practices.

The claim here is that mental logic can contribute towards the fulfilment of these conditions in support of the language of thought hypothesis. It has, in fact, been the burden of the works of mental logicians to include analyses of components and errors which permit one to predict where interaction effects should be expected and thereby allow justified appeals to the competence/performance distinction in a *principled* way. They undertake to show that so many apparent violations of *rational* inferential behavior can be explained away without resorting to dirty tricks. Indeed, in Braine, *et al.*'s proposed program for propositional reasoning, one can specify subjects' logical competence and state the *minimal closure condition* along the following form: propositional reasoning is closed under application to *primary reasoning skills* (Braine and O'Brien 1998).[9]

---

[8]O'Brien (2010), for example, extends the mental logic theory to another fragment of reasoning dealing with conditional propositions.

[9]'Primary reasoning skills' is the title that Braine bestows on his theory of mental logic which is cardinally concerned with the "information integration processes of a logical sort that we share with the rest of humanity, literate and illiterate" (1990, 137). These primary routines stand in contrast to *secondary* logical reasoning skills that Braine classifies as *quasi-academic* which depend on literacy and a concern with language *qua* language, that is, such skills that require a degree of compartmentalization of information

What is significant about this condition is that it allows one to specify which failures of classically acceptable inferential patterns depend on subjects' competence limitations. For example, one often refrains from accepting 'not A' even if a contradiction can be derived from 'A'. This is because unconstrained *reductio ad absurdum* strategies are not accessible for most subjects in contrast to the availability of the principle 'from a contradiction nothing follows.' Also, the minimal closure condition seems able to specify which substitutions of logically identical sentences – or, to use the syntactic mode of speech, which syntactic transformations – are allowed *within* the subjects' cognitive space and which ones are ruled out. Similarly, within its own limits, the primary reasoning skills theory is also capable of predicting observed inferential practices which constituted the third thread in the original question, *viz.* to ask for a theory with adequate predictive power.

Insofar as the problem of opacity is concerned, one may hypothesize that once the closure conditions for propositional reasoning are set, the theory would also be able to set the limits of opaqueness in intentional contexts. This, in turn, may shed light on how the language of thought hypothesis can avoid excessive logical transparency, i.e. the problem of logical omniscience, without introducing excessive opaqueness, *viz.* the problem of the poverty of laws. This means that subjects are not logically omniscient because their inferential system is weaker than classical logic, and, thus, not even in point of competence should one attribute to them knowledge of either all logical truths or all the logical consequences of their held beliefs.

Yet, one is not totally blind to people's systems of beliefs. One could rather safely assume that, *qua* humans, if they have a belief they will also have other occurrent or dispositional beliefs: minimally, all those that can be inferred on the basis of the innate logic in the mind. For example, if Adam believes that P and Q, then one is entitled to ascribe to him, certainly *dispositionally* if not *occurrently*, the belief that P and the belief that Q on the ground of a logical law of elimination of conjunction which applies in intentional contexts and should be part of Adam's reasoning competence. Similarly, if he believes that P and that P implies Q, and that Q implies not P, then by transparent *modus ponens* and by the principle that 'from a contradiction nothing follows' one is licensed to anticipate that Adam will start a process of revision of his beliefs. Certainly, one is not going to expect Adam to believe anything whatsoever including contradictory propositions, since *ex falso quodlibet* is not a theorem of natural logic, even if the procedure happens to be classically valid.[10]

Therefore, it seems that mental logic is able to give the language of thought hypothesis the empirical basis it needs to show that it can deal with the issues it was devised for in the first place. Generally, mental logic's liaison with the hypothesis seems to secure some degree of progress for a constellation of problems in the psychophilosophy of cognition. Yet, there remains one question: although the language of thought hypothesis clearly benefits from its liaison with mental logic and the latter is able to give sustenance to it in various ways, how secure is the hypothesis on its own merits? Perhaps one should leave the last word on this matter to Fodor himself: "[w]hich is not to deny that there are [ahem!] certain residual technical difficulties" (1988, 156).

---

and analytic comprehension which are a matter of acquisition.

[10]This is the *principle of explosion* that 'from falsehood anything (follows)'.

Works Consulted

Braine, M. D. "On the Relation Between the Natural Logic of Reasoning and Standard Logic." *Psychological Review* 85 (1978): 1-21.

Braine, M. D. "The 'Natural Logic' Approach to Reasoning." *Reasoning, Necessity, and Logic: Developmental Perspectives*. Ed. W. Overton. Hillsdale, NJ: Lawrence Erlbaum, 1990. 133-158.

Braine, M. D. "Mental Logic and How to Discover It." *The Logical Foundations of Cognition*. Ed. J. Macnamara and G. E. Reyes. Oxford: OUP, 1994. 241-265.

Braine, M. D., and D. P. O'Brien, eds. *Mental Logic*. Mahwah, NJ: Lawrence Erlbaum, 1998.

Braine, M. D., B. J. Reiser, and B. Rumain. "Some Empirical Justification for a Theory of Natural Propositional Logic." *The Psychology of Learning and Motivation* 18 (1984): 313-371.

Braine, M. D., and B. Rumain. "Logical Reasoning." *Handbook of Child Psychology: Cognitive Development*. Vol. 3. Ed. J. Flavell and E. Markman. Chichester: John Wiley & Sons, 1983. 263-340.

Carnap, R. *Meaning and Necessity*. 2nd Edition. Chicago: U of Chicago P, 1988.

Davies, M. "Concepts, Connectionism, and the Language of Thought." *Philosophy and Connectionist Theory*. Ed. W. Ramsey, S. Stich, and D. Rumelhart. Hillsdale, NJ: Lawrence Erlbaum, 1991. 229-258.

Dowty, D. R., R. E. Wall, and S. Peters. *Introduction to Montague Semantics*. Dordrecht: D. Reidel, 1981.

Fodor, J. A. *The Language of Thought*. Brighton: Harvester Press, 1976.

Fodor, J. A. *Representations*. Brighton: Harvester Press, 1981.

Fodor, J. A. *Psychosemantics*. Cambridge, MA: MIT Press, 1988.

Fodor, J. A. *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press, 1990.

Fodor, J. A. *LOT 2: The Language of Thought Revisited*. Oxford: Clarendon, 2008.

Henle, M. "On the Relation Between Logic and Thinking." *Psychological Review* 69 (1962):366-378.

Kripke, S. "A Puzzle About Belief." *Meaning and Use*. Ed. A. Margalit. Dordrecht: D. Reidel, 1979. 239-283.

Lewis, D. "General Semantics." *Semantics of Natural Language*. 2nd Edition. Ed. D. Davidson and G. Harman. Dordrecht: D. Reidel, 1972. 169-218.

Lycan, W. "A Deductive Argument for the Representational Theory of Thinking." *Mind and Language* 8 (1993): 404-22.

Matthews, R. "The Alleged Evidence for Representationalism." *Representation*. Ed. S. Silvers. Dordrecht: Kluwer Academic, 1989. 103-122.

Matthews, R. "Is There Vindication Through Representationalism?" *Meaning in Mind: Fodor and His Critics*. Ed. B. Loewer and G. Rey. Oxford: Blackwell, 1991. 137-149.

Matthews, R. "The Measure of Mind." *Mind* 103 (1994): 131-146.

May, R. *Logical Form: Its Structure and Derivation*. Cambridge, MA: MIT Press, 1985.

May, R. "Logical Form as a Level of Linguistic Representation." *New Directions in Semantics*. Ed. E. LePore. London: Academic Press, 1987. 305-336.

O'Brien, D. P., and A. Manfrinati. "The Mental Logic Theory of Conditional Propositions." *Cognition and Conditionals: Probability and Logic in Human Thinking*. Ed. M. Oaksford and N. Chater. Oxford: OUP, 2010. 39-54.

Osherson, D. "Logic and Models of Logical Thinking." *Reasoning: Representation and Process in Children and Adults*. Ed. R. J. Falmagne. Hillsdale, NJ: Lawrence Erlbaum, 1975. 81-92.

Osherson, D. "Natural Connectives: a Chomskyan Approach." *Journal of Mathematical Psychology* 16 (1977): 1-29.

Rey, G. "'A Not "Merely Empirical' Argument for a Language of Thought." *Philosophical Perspectives* 9 (1995): 201-222.

Rips, L. "Cognitive Processes in Propositional Reasoning." *Psychological Review* 90 (1983): 38-71.

Rips, Lance. *The Psychology of Proof*. Cambridge, MA: MIT Press, 1994.

Sober, E. "The Evolution of Rationality." *Synthese* 46 (1981): 95-120.