RIFL/SFL(2015): 78-91
DOI: 10.4396/2015SFL07

Is novelty of thought a mystery? Old schemata for new problems*

Marco Mazzone
Dipartimento di Scienze Umanistiche, Università di Catania
mazzonem@unict.it

Abstract Contextual flexibility has been the focus of considerable research in cognitive science. The pessimistic view expressed by Fodor on this issue has been challenged either by modular approaches or by proposals based on common codes/spaces where information can be integrated. I analyse these views and the middle-ground approach explored by Shanahan and Baars (2005), and propose a different non-modular account. The general idea is that flexible integration of information is essentially ensured thanks to the hierarchical and schematic organization of memory, and to the bottom-up/top-down dynamic of its associative activation. In this perspective, context is not a crucial part of the problem, it is instead key to the solution: the different inputs in a context activates schemata that compete and integrate with each other, so that the schemata that are the most coherent with the context will be the most activated as well. I also consider the role that might be played by consciousness in this process, especially with regard to cases of extreme flexibility, that is, cases in which creative thoughts are formed.

Keywords: cognitive flexibility, schemata, consciousness, frame problem

Received 11 February 2016; received in revised form 01 July 2016; accepted 05 July 2016.

0. Introduction

In the last decades, the issue of contextual flexibility in human behaviour has received more than passing attention in cognitive science. From a philosophical perspective, Fodor (2000) has notoriously «developed the dark message of *The Modularity of Mind* (FODOR 1983)» (WILSON 2008: 407), by insisting that while peripheral and modular processes are computationally tractable, central processes – precisely because of their flexibility – are not, and therefore they are bound to remain a mystery for cognitive science. Fodor's challenge, however, has been taken up by some scholars convinced that contextual flexibility is not a mystery, it is instead a problem that can and must be addressed.

In this paper, I want to consider the extreme case of contextual flexibility, that is, the case of thoughts that appear genuinely novel – as when, for instance, we devise a

* I want to thank two anonymous referees of RIFL and the friends who participated to the XXII Conference of SIFL (Società italiana di filosofia del linguaggio) for comments on previous versions of this work.

creative solution for a problem. My aim is to suggest that even in such a case flexibility is substantially dependent on the structure of memory and therefore, in a sense, novelty leans on old information. It is important to emphasize that my purpose is not to take an eliminativist stance towards novelty of thought[1], nor to pursue an evocative paradox for the sake of paradox. There is a clear sense in which genuine novelty is, after all, novel, not old, and in fact I will point out some crucial differences between mere reactivation of schemata already in our repertoires and the formation of new thoughts. However, my claim is that the organization of memory – of old information – is key to making flexibility far less mysterious than it might seem, and that this has something interesting to say also about extreme cases of flexibility.

In practice, I will shortly examine the main solutions to Fodor's problem offered in the literature, arguing that they fall short of providing a satisfactory account of flexibility. I will then defend the thesis that flexibility is essentially dependent on the structure of long-term memory, and specifically on its schematic and hierarchical organization. This organization allows a bottom-up/top-down dynamic that is not encapsulated but distributed instead across different domains. I will also consider the role that might be played within this picture by working memory and consciousness. In a sense, my whole proposal can be seen as a mediation between the accounts based on automatic processes working in parallel (e.g., SPERBER 2005) and the ones based on common spaces or codes (e.g., CARRUTHERS 2003; 2006). As such, my view is similar to the one defended by Shanahan and Baars (2005), except that they do not see the contradiction that is inherent in appealing to modular processes in order to solve the problem of encapsulation. As a consequence, they are not very clear about the role played by consciousness as well: consciousness can actually be described as a common space for the exchange and comparison of information, but this is not something that adds to encapsulated modular processes operating in parallel, it is instead something that modulates automatic parallel processes and makes them even less modular than they would be anyway.

While all of this applies to contextual flexibility in general, I will then analyse what is specific of genuine cases of novel thoughts. In particular, I will consider again the role of consciousness as a possible explanation, and I will suggest that this cannot work, at least in the most obvious interpretation. Consciousness seems to operate in favour of stability and persistence – not flexibility – in processing (HOMMEL 2015). On the other hand, however, persistence can be indirectly exploited in favour of flexibility, by actively focussing on problems and questions, so as to inhibit the most obvious and automatic solutions/answers and leave room for active search of different ones. While questions by themselves constrain the space of the search to some extent, it is often crucial that some more specific schema is employed to drive the search towards appropriate targets. The intuitive idea I want to defend is that no novel solution can be found and recognized *as* the solution to a given problem unless one gets some hints about what she is looking for. A solution is novel when there is no schema that *directly* provides it: but unless one has some schema, abstract and indirect as it can be, which is apt to suggest the solution, no solution can be found – or even recognized – at all. Thus, I intend to take seriously the suggestion made by Herbert Simon, when he claims that smart intuition is, in the end, just recognition:

[1] I am indebted to Paolo Leonardi for having drawn my attention to the need of this warning.

> The situation has provided a clue; this clue has given the expert access to information stored in memory, and the information provides the answer. Intuition is nothing more and nothing less than recognition (SIMON 1992; quoted by KAHNEMAN 2011: 11).

1. The frame problem and possible solutions

According to Fodor (1983; 1987; 2000), the dire problem affecting flexible central processes is in essence what McCarthy and Hayes (1969) called the frame problem in Artificial Intelligence and Dennet (1984) famously restated in more evocative – but also vague – terms. The point is that unencapsulated processes appear to be computationally intractable for a finite mind.

> The totality of one's epistemic commitments is vastly too large a space to have to search if all one's trying to do is figure out whether, since there are clouds, it would be wise to carry an umbrella. Indeed, the totality of one's epistemic commitments is vastly too large a space to have to search *whatever* it is that one is trying to figure out (FODOR 2000: 31).

To be sure, doubts have been raised about whether the notion of computational intractability is used properly by Fodor and whether the original computational problem is anyway as damaging as it may appear. Shanahan and Baars (2005) argue that both questions deserve a negative answer. In particular, they observe that in computer science individual instances of intractable problems «are frequently soluble in reasonable time» (SHANAHAN, BAARS 2005: 162) and that there exist a variety of practical techniques to mitigate the effects of intractability. However, they also observe that the intuitive force of the argument cannot be denied: «The task remains to explain how an informationally unencapsulated process might be realized in a biological brain» (*Ivi*: 164). More precisely, the problem to be addressed is how biological brains manage to sift the relevant from the irrelevant for any cognitive task.

One possible option is to deny that unencapsulated processes are needed at all in order to get flexibility and contextual selection of (the most) relevant information. This is the option chosen, for instance, by Sperber (2005). In his view, minds are characterized by massive or even teeming modularity, that is, they can be described as nothing else than a variety of modules competing and cooperating with one another. Now, Sperber's idea is that the competition between modules is itself the way in which our cognitive systems, as a result of evolutionary pressure, have solved the problem of sifting the relevant from the irrelevant, without any need for central unencapsulated processes. The attentional bottleneck can be expected to create in fact

> a strong and constant selective pressure for optimizing the choice of inputs to be processed […]. Such a selective pressure should result in the evolution of a variety of traits contributing to an optimal allocation. […] I find it implausible, both for evolutionary and efficiency reasons, to imagine that this allocation of resources might be wholly or even mostly controlled by some central specialized device (SPERBER 2005: 66).

I cannot go into the details of Sperber's argument. I just note that the idea that evolution might have solved the problem for us, as intuitively appealing as it may be, leaves essentially unaddressed the question of how module encapsulation may allow flexible integration of information from different domains. Even supposing that modules compete with each other for the processing of their own data, it is also required that they integrate those data in a way that changes as a function of infinitely variable contexts. But this seems at odds with the very intuition underlying the notion of *encapsulated* modules, whose main point is precisely that they *do not* exchange information with each other flexibly[2].

On the opposite side of the debate, some have proposed that cognitive flexibility requires a common space and/or a common code in which information coming from a variety of sources can be confronted and integrated. A case in point is Carruthers (2003; 2006). In his proposal, «natural language is the medium for non-domain specific thinking, serving to integrate the outputs of a variety of domain-specific conceptual faculties» (CARRUTHERS 2003: 657). However, it is far from clear that the existence of a common space/code for the integration of cross-domain information is sufficient to ensure flexibility. For instance, as Machery (2008) noted, since Carruthers adopts the view that conceptual processes are modular, and since conceptual modules are held to process only contents that are specific for their domains, one wonders which module should process contents resulting from cross-domain integration. But proposals of that kind also face a problem that is, as far as I can tell, even worse than that. Integration is only one side of the problem; the other side, as we saw, is selection of relevant information. Thus, one has to wonder how the appropriate information is selected for integration. Either it is already selected by modules and their competition, and then we are back to the problem discussed above, or all the information inside the system is virtually made available to the common space/code, but then there needs to be a central unencapsulated mechanism that considers this huge amount of information and picks out the most relevant bits, which is precisely the kind of mechanism that Fodor warned us to avoid.

Here the point is that if modules are conceived of as passive transmitters of information, the problem of selecting relevant information can only be solved by central mechanisms of sort. This is why Shanahan and Baars focus on the need for «multiple, parallel processes that all contribute *actively* to cognition» (SHANAHAN, BAARS 2005: 168; their emphasis). In their view, this is compatible with their proposal of a «global workspace architecture» whose essence is

> a model of combined serial and parallel information flow. Multiple parallel *specialist processes* compete and co-operate for access to a *global workspace*. […] If granted access to the global workspace, the information a process has to offer is 'broadcast' back to the entire set of specialists. […] The means by which access is granted to the global workspace can be likened to an attention mechanism (*Ivi*: 165; their emphasis).

There is much to be recommended in this approach. In particular, it is consistent with the fact that automatic parallel processes and conscious serial ones cooperate with each other in most cognitive tasks, and that attention has a bottom-up component, based on competition between information processed in parallel, and a top-down one, thanks to which attended information is made available for further processing. However, a couple of remarks are in order.

2 A wider discussion of Sperber's and Carruther's accounts for flexibility can be found in MAZZONE (2015a: 27-8, 48-51, 96-104).

RIFL/SFL(2015): 78-91
DOI: 10.4396/2015SFL07

In the first place, Shanahan and Baars claim that «Global workspace theory, with its commitment to multiple, parallel, specialist processes, is consistent with a *modular view* of the human mind» (*Ivi*: 174; emphasis mine), insofar as «the responsibility for deciding relevance [is distributed] to the parallel specialists themselves» (*Ibidem*; emphasis mine). But if this is taken in a reasonably strict sense of *modular*[3], we are back to the problem of how encapsulated modules may ensure flexible integration of information across domains.

One possibility is that the problem is overcome precisely by combining the two views. At first sight, while the modular view seems able to account for selection (*via* competition) but not for integration of information, the central workspace view shows the opposite pattern: it provides a virtual space for integration, without any account of how the to-be-integrated information is selected. Thus, one might hope that modular processing ensures the selection of relevant information, while its integration is accomplished in the central workspace. However, it is disputable that there can be a mechanism for the selection of *relevant* information that is blind to its changeable integration. Information is either relevant or not *with regard to contexts*. But as we noted, *encapsulated* modules do not – by definition – exchange information flexibly in the course of processing and therefore they can only blindly compete with each other. As a consequence, it is left underspecified, at the very least, how each module might convey to the central workspace the contextually relevant bits of information. In the same vein, we should keep in mind Machery's criticism: once information is integrated how should it be further processed, given that conceptual modules are domain-specific?

In sum, selecting the appropriate information seems to require unencapsulated and contextually sensitive access to it (just as processing integrated information seems to require non-domain-specificity). In line with these suggestions, in the next section I will propose a different attempt to reconcile the modular (or rather, the parallel processing) view and the central space view than the one proposed by Shanahan and Baars. The model I propose capitalizes on the idea that the responsibility for deciding relevance is highly distributed, but within a non-modular framework which allows for the integration of information at any point in processing. In practice, selection of information is proposed to be the result of a distributed process of integration. Moreover, I will argue that the selection of relevant information is largely dependent on automatic parallel processing, but with consciousness playing a role of its own.

2. Contextual flexibility and the organization of memory

Associative memory is not a passive repository for information. On the contrary, the way in which information is coded in memory plays a crucial role in its subsequent recovery and processing. In particular, as noted by Sperber and Wilson, «memory is so organized that pieces of information that are likely to be *simultaneously relevant* tend to be co-accessed or co-activated in chunks variously described in the literature as 'concepts', 'schemas', 'scripts', dossiers', etc.» (SPERBER, WILSON 1996: 531; emphasis mine). By developing this intuition, an entirely different light is shed on the issue of flexibility: the search for relevance can be seen as driven by the organization of memory, and especially by its schematic and hierarchical structure. Due to this

[3] In MAZZONE (2015c), I provided a detailed argument to the effect that modularity (and domain-specificity in particular) is tenable only in a weak sense, which makes the modularity thesis trivial.

structure, the context, far from being at the core of the problem of flexibility, is key to its solution.

Let us first sketch the general background. In his insightful synthesis of the current knowledge about human cortex, Fuster (2003) described it as a unitary memory space, highly interconnected and hierarchically organized. In particular, interconnectivity between domains can be presumed to be high as a consequence of the Hebbian principle governing the cortex: neural patterns that are repeatedly co-activated are associated with each other, irrespective of whether the related inputs belong to the same domain or not.

In the same vein, in theory of concepts Barsalou (e.g., 2005) defended the view that concepts are *situated* representations, meaning that they preserve contextual information about specific settings in which objects are usually located. Yeh and Barsalou provide a review of the evidence showing that concepts are situated, and that this affects «conceptual processing in many tasks [...] across many areas of cognition» (BARSALOU, YEH 2006: 349). It is well established that the recovery of information from memory shows contextual effects even for non-relevant contextual features: for instance, what is learned in one environment (or in a language) is better recalled in it than in alternative settings (or languages) (e.g., GODDEN, BADDELEY1980).

The general lesson is that we detect and store in memory regular patterns of covariation, both within and *between* domains, with the consequence that the recovery of information based on associations is not confined within domain boundaries. Barsalou described this mechanism – with regard to what he calls «conceptualizations», that is, wider structures of information in which concepts are embedded – in the following terms:

> The conceptualization is essentially a pattern, namely, a complex configuration of multimodal components that represent the situation. When a component of this pattern matched the situation, the larger pattern became active in memory. The remaining pattern components – not yet observed – constitute inferences, that is, educated guesses about what might occur next (BARSALOU 2005: 628).

There are here three points that deserve further consideration. First, despite its simplicity the mechanism described by Barsalou is powerful enough to provide a robust and general account of parallel processing, which is alternative to the modularity thesis. Second, in contrast with the modularity thesis, the alternative view can easily account for contextual flexibility. Third, this view can easily account for the cooperation between automatic parallel processing and conscious processing. Let us address these points in turn.

In the previous quotation, Barsalou describes a mechanism of bottom-up/top-down activation mediated by what I will call *schemata*, that is, complex configurations of components[4]. The activation of one component of the schema spreads to the whole schema (bottom-up activation) and this activates in turn its other components (top-down activation). Assuming an extensive hierarchical organization of information, such a mechanism is powerful enough to explain generative processing without any specialized process. Perhaps the clearest illustration of this point is the framework

[4] The notion is analysed at some length in MAZZONE (2014a); MAZZONE (2015b) provides a wider analysis of the bottom-up/top-down dynamic.

proposed by Jackendoff (2007a) for syntax. He observes that while Generative Grammar adopts procedural rules in order to account for the fact that we can produce an infinite number of sentences with a finite set of linguistic units, the same result can be gained by substituting encoded schemata for procedural rules. In other words, schemata can provide exactly the same information that procedural rules were supposed to, insofar as the former are higher-level structures prescribing the organization of lower-level components. An important feature of this approach is that it has no need to postulate a set of specialized procedures (a module) for each domain. Jackendoff (2007a) proposes, in fact, that a single general-domain process, which he calls *unification*, may account for combinatorial processes across different cognitive domains – a natural interpretation of this process being in terms of the simple bottom-up/top-down activation described above. In this framework, the within-domain organization of information is sufficient to ensure domain-specificity of processing; however, since between-domain relationships are also coded, the resulting pattern of activation is far from being one of rigid encapsulation[5].

The versatility of the described mechanism is shown by its wide application to other domains. In particular, the idea of a generative mechanism based on a hierarchical organization of information is largely explored in theories of action and motor control (e.g., COOPER, SHALLICE 2006; GLENBERG, GALLESE 2012; JACKENDOFF 2007b; MAZZONE 2014b; 2014c; PASTRA, ALOIMONOS 2012). Moreover, there is neurocomputational and neuroscientific literature suggesting the existence of representational hierarchies in the brain, especially with regard to frontal areas (BOTVINICK 2008; FUSTER 2001; 2003; GRAFTON, HAMILTON 2007)[6].

The mechanism of bottom-up/top-down activation mediated by schemata not only provides an account of generative processing that is alternative to the modular view, but it also allows accounting for contextual flexibility. Let me show this with an example drawn from the domain of pragmatic understanding. Recanati (2004) explained how such a mechanism may have the effect of increasing coherence in utterance understanding. For instance, given the utterance «The ATM swallowed the credit card», the concept *swallow* is modulated in context by selecting the features that are coherent with the other concepts, ATM and *credit card*. This process of selection can be described in the following terms. While the concept *swallow* activates a number of associated schemata, including the ones that activate in turn its constitutive properties, the other concepts activate their associated schemata as well.

[5] For a recent development of the hierarchical model for language processing, see BROWN, KUPERBERG (2015). An anonymous referee has expressed doubts as to whether the combinatorial aspects of language generation (both in syntax and semantics) are reducible to hierarchies based on schemata. The reducibility thesis is defended by Jackendoff, who claims that schematic representations are able to explain the same data than procedural rules and others as well, insofar as the former (but not the latter) allow for violation and competition between constraints (JACKENDOFF 2007a: 9; JACKENDOFF 2007c: 358; see also MAZZONE 2011: 2156). The key idea is that you can have a schema for any of the procedural rules posited by Generative Grammar, with the former containing exactly the same information than the latter, but the schema-based account is not bound to host only compatible rules and thus it can be enriched so as to explain conflicts between rules. The reducibility of generativity to hierarchical organization is either explicitly claimed or implicitly presupposed by the literature mentioned in the next paragraph, and I have especially focussed on this issue in MAZZONE (2014a, 2014b, 2014c, 2015b).

[6] DUNCAN (e.g., 2010) appeals to the hierarchical organization of information in order to account for what he calls «fluid intelligence», that is, flexibility of thought in a slightly different sense: the ability to change flexibly our strategy in online tasks.

To the extent that these latter schemata have as their components the features also activated by *swallow*, those features receive further activation and therefore – so to speak – win the competition. More generally, any input (here, words and the related concepts) can be expected to activate the associated schemata, thus providing a variety of expectations about the surrounding inputs. Any of those expectations, on the other hand, can either correctly predict the remaining inputs or not, thus receiving or not further activation from them. As a result, the most activated schemata are also the most coherent with the situational context, and they in turn set the interpretation of the inputs.

There is a lesson to be drawn from these considerations. We started by describing the problem of flexibility as the problem of how the information that is relevant in context is selected. Under that description, the infinite variability of contexts seemed to lie at the core of the problem: how is it that the appropriate information is selected on any occasion, in response to always changing contexts? However, in the light of our previous suggestions, the situational context appears instead as part of the solution. The relevant features of the concept *swallow*, for instance, are selected thanks to the fact that its context (specifically, the other concepts expressed by the utterance) provides constraints on the appropriate interpretation. In sum, contextual flexibility is nothing but the selection of information activated, through integration and competition, *by* the context at hand.

As a final point, I want to consider the relationship between automatic parallel and conscious serial processing. In the model proposed by Shanahan and Baars (2005) the cooperation between the two can be thought of as a straightforward division of labour, in which the conscious global workspace is responsible for the integration of the information selected by multiple parallel specialist processes. In contrast, I have defended the view that information is integrated incrementally right from the beginning. If this is the case, one might wonder whether integration is entirely performed automatically and in parallel, without any role to play for consciousness. A far more plausible picture, though, is the following (see also MAZZONE, CAMPISI 2013; MAZZONE 2013a; 2013b; 2016; in press). Automatic and conscious processes cooperate with each other in any cognitive task of reasonable complexity. Specifically, there is a continuous flow of activation proceeding automatically and in parallel, a small part of which gains access to consciousness. According to a widely accepted model (see for instance DEHAENE 2014; DEHAENE et al. 2006)

> while automatic processes consist in weak and local spreading activation with fast rise time and rapid decay time, conscious processes consist instead in strong activation of distant areas in the cortex (typically, both frontal and posterior sensorimotor areas), which form long-distance loops that are self-sustaining and thus enduring in time […]. In other words, consciousness is essentially a mechanism for the maintenance of activation, which makes it possible even for representations with very indirect connections to act on each other (MAZZONE, in press).

For this long-distance integration to occur, however, it is not necessary that each of the representations involved, and their relation, be attended consciously. Consciousness can also occur in a «low-effort mode» (KAHNEMAN 2011: 24) thanks to which, for the simple fact that conscious attention focuses on some of the related representations, the others are non-consciously activated.

To sum up, what I have proposed is a domain-general mechanism – based on bottom-up/top-down associative activation mediated by schemata – for the integration of (and competition between) information, which also accounts for its flexible contextual selection. Not only is this mechanism domain-general, it also depends on the structure of associative memory, which is likely to be hierarchically organized but far from encapsulated: it preserves instead a rich network of cross-domain connections. Finally, conscious attention and automatic processing cooperate in making that integration possible, thanks to the fact that the latter allows parallel processing while the former accounts for longer-range, time-extended interactions.

3. Novel thoughts and consciousness

With all this in mind, we can now turn to what I have called extreme flexibility, that is, the ability to form genuinely novel thoughts. One may be tempted to think that extreme flexibility depends essentially on conscious reasoning, although this intuition might follow from a general tendency to overestimate the role played by conscious control in our lives (e.g., WILSON 2002). As a matter of fact, I have previously characterized consciousness as a mechanism for the maintenance of activation, and in the same vein Hommel (2015) has recently insisted that cognitive control can be seen as a dynamic interaction between persistence and flexibility, with persistence, not flexibility, being ensured by the brain areas typically associated with conscious control (prefrontal cortex) – while the striatum would be responsible for flexibility instead.

Persistence (maintenance of activation) is especially important for goal-directed actions, since these require that the relevant information remain active until the goal is achieved. However, persistence is instrumental in other cognitive functions as well, especially in the monitoring of ongoing activities. To this purpose, the previously introduced notion of schema appears to be crucial: conscious monitoring of cognitive activities implies that a schematic representation of the task at hand is maintained active, so that the system can check whether the different components of the task are aptly accomplished.

The importance of this mechanism is made apparent by considering a well-known bias of our cognitive system: «If a satisfactory answer to a hard question is not found quickly, System 1 [i.e., the mechanism for automatic associative processing] will find a related question that is easier and will answer it» (KAHNEMAN 2011: 97). This is a consequence of the «spreading cascade of activity» (*Ivi*: 51) that characterizes associative activation, which in turn is responsible for what Kahneman calls the

> mental shotgun: we often compute much more than we want or need […]. It is impossible to aim at a single point with a shotgun because it shoots pellets that scatter, and it seems almost equally difficult for System 1 not to do more than System 2 [i.e., the control mechanism] charges it to do (*Ivi*: 95).

In normal circumstances, conscious attention is in a comfortable low-effort mode, and is prone to adopt the suggestions of automatic processes with no modification (*Ivi*: 24). When required, however, we can abandon the low-effort mode and pay special attention to the answers provided by automatic activation, so as to check them against the original questions or tasks. To that purpose, we must keep active the task representation, together with further schemata specifying the inferential procedures (demonstrative or not) that are relevant to it. Such inferential schemata, I submit, might have different degrees of abstraction, from domain-general logic rules to domain-specific regularities (possibly including memories of individual instances), as a function of personal experience with formal reasoning and/or more specific bodies of knowledge. This variety of schemata at different degrees of abstraction might explain the different patterns of results in studies of reasoning, some favouring formal rules (HENLE 1962), others favouring what Cheng and Holyoak (1985) called «pragmatic schemas» based on domain-specific knowledge.

Now, what happens when, in the presence of such a conscious monitoring, «System 1» cannot provide any answer or, at least, any answer that is consistent with the demands of the task? First of all, we can expect that inappropriate answers be rejected, which amounts to their active inhibition. At this point, there are different options. We may launch an active search in our associative memory, that is, we adopt a relatively effortful conscious mode allowing the activation, and recovery, of information only indirectly related to our question/task. Or we may keep the question in our short-term memory and wait until we run into information that is helpful in solving the problem.

In any case, the representation of the task provides constraints on the solution, but often not enough to guide our search in the right direction. This is why analogy has been proposed to be a key resource for reasoning: it may provide schemata that are apt to (re)frame with sufficient specificity, and thus solve, a problem. One might even go further and suggest that analogical reasoning presupposes a general *schema for constructing analogies*, that is, a schema for employing schemata from source domains in order to highlight structural similarities with target domains.

Then, what might be the role of consciousness in analogical reasoning? Fodor observed that reasoning by analogy is a typical case of creative, unencapsulated cognitive process because it «depends precisely upon the transfer of information among cognitive domains previously assumed to be irrelevant» (FODOR 1983: 105). So, Shanahan and Baars (2005) feel it important to show that this ability can be accounted for in their perspective. In practice, they analyse the most effective computational models of reasoning by analogy and point out that their success is essentially dependent on retrieval being cast as a parallel process and mapping as inherently serial. Based on our previous considerations, the suggestion that the search for information in memory is largely a parallel automatic process does make sense. A «spreading cascade of activity», in Kahneman's words, is the basic mode of operation for associative memory. It is debatable, however, that serial conscious processing has no role to play in guiding the search. In particular, when we have to draw creative analogies, as these involve the activation of schemata not associated to the target domain in any direct way, the search is likely to require the kind of strong and persistent activation that is a distinctive feature of conscious attention.

To sum up, my proposal is that extreme flexibility is a relatively effortful process requiring the contribution of conscious attention. Specifically, it requires that the task at hand be consciously focused so that an active search for distant information is launched. For instance, in reasoning by analogy the target domain has to be consciously attended so that its inherent structure is somehow appreciated and the search for similar schemata is launched. However, another thesis that I have defended is that no problem can be solved or cognitive task be accomplished unless we find in our memory[7] the instruction on how the appropriate information must be

[7] Or in the world, of course: but that would not count as a case of flexibility in mental operations.

integrated. To be sure, this is quite different from the automatic integration of information that is directly associated. But, as I hope to have argued convincingly, the solution of a problem cannot even be recognized *as* the solution, let alone be recovered from memory, in the absence of schemata driving the search – schemata that tell us, in fact, how the solution might look like.

References

BARSALOU, Lawrence W. (2005), *Situated conceptualization*, in COHEN Henry, LEFEBVRE Claire (a cura), *Handbook of Categorization in Cognitive Science*, Elsevier, St. Louis, pp. 619-650.

BOTVINICK, Matthew Michael (2008), «Hierarchical models of behavior and prefrontal function», in *Trends in Cognitive Sciences*, n. 12 (5), pp. 201-208.

BROWN, Meredith, KUPERBERG, Gina (2015), «A Hierarchical Generative Framework of Language Processing: Linking Language Perception, Interpretation, and Production Abnormalities in Schizophrenia», in *Frontiers in Human Neuroscience*, n. 9 (643), pp. 1-23.

CARRUTHERS, Peter (2003), «On Fodor's problem», in *Mind & Language*, n. 18, pp. 502-523.

CARRUTHERS, Peter (2006), *The Architecture of the Mind: massive modularity and the flexibility of thought*, Oxford University Press, Oxford.

CHENG, Patricia W., HOLYOAK, Keith J. (1985), «Pragmatic reasoning schemas», in *Cognitive Psychology*, n. 17, pp. 391-416.

COOPER, Rick, SHALLICE, Tim (2006), «Hierarchical schemas and goals in the control of sequential behavior», in *Psychological Review*, n. 113 (4), pp. 887-916.

DEHAENE, Stanislas (2014), *Consciousness and the Brain: Deciphering how the Brain Codes our Thoughts*, Viking, New York.

DEHAENE, Stanislas, CHANGEUX, Jean-Pierre, NACCACHE, Lionel, SACKUR Jérôme, SERGENT Claire (2006), «Conscious, preconscious, and subliminal processing: a testable taxonomy», in *Trends in Cognitive Sciences*, n. 10 (5), pp. 204-11.

DENNETT, Daniel (1984), *Cognitive wheels: The frame problem in artificial intelligence*, in HOOKWAY Cristopher (a cura), *Minds, machines and evolution*, Cambridge University Press, Cambridge, pp. 129-151.

DUNCAN, John (2010), «The multiple-demand (MD) system of the primate brain:

mental programs for intelligent behaviour», in *Trends in Cognitive Sciences*, n. 14(4), pp. 172-179.

FODOR, Jerry Alan (1983), *The modularity of mind*, MIT Press Cambridge (Massachusetts).

FODOR, Jerry Alan (1987), *Modules, frames, fridgeons, sleeping dogs, and the music of the spheres*, in PYLYSHYN Zenon Walter (a cura), *The robot's dilemma: The frame problem in artificial intelligence*, Ablex, Norwood (New Jersey), pp. 139-149.

FODOR, Jerry Alan (2000), *The mind doesn't work that way*, MIT Press, Cambridge (Massachusetts).

FUSTER, Joaquin (2001), «The prefrontal cortex – An update: Time is of the essence», in *Neuron*, n. 30(2), pp. 319-333.

FUSTER, Joaquin (2003), *Cortex and Mind*, Oxford University Press, Oxford.

GLENBERG, Arthur, GALLESE, Vittorio (2012), «Action-based language: A theory of language acquisition, comprehension and production», in *Cortex*, n. 48, pp. 905-922.

GODDEN, Duncan, BADDELEY, Alan (1980), «When does context influence recognition memory?», in *British Journal of Psychology*, n. 71(1), pp. 99-104.

GRAFTON, Scott, HAMILTON, Antonia (2007), «Evidence for a distributed hierarchy of action representation in the brain», in *Human Movement Science*, n. 26, pp. 590-616.

HENLE, Mary (1962), «On the relation between logic and thinking», in *Psychological Review*, n. 69, pp. 366-378.

HOMMEL, Bernard (2015), *Between persistence and flexibility: The Yin and Yang of action control*, in ELLIOTT Andrew (a cura), *Advances in motivation science, Vol. 2*, Elsevier, New York, pp. 33-67.

JACKENDOFF, Ray (2007a), «A parallel architecture perspective on language processing», in *Brain Research*, n. 1146, pp. 2-22.

JACKENDOFF, Ray (2007b), *Language, Consciousness, Culture: Essays on Mental Structure*, MIT Press, Cambridge (Massachusetts).

JACKENDOFF, Ray (2007c), «Linguistics in cognitive science: the state of the art», in *The Linguistic Review*, n. 22, pp. 347-401.

KAHNEMAN, Daniel (2011), *Thinking, Fast and Slow*, Farrar, Straus and Giroux, New York.

MACHERY, Eduard (2008), «Massive modularity and the flexibility of human cognition», in *Mind & Language*, n. 23(3), pp. 263-272.

MAZZONE, Marco (2011), «Schemata and associative processes in pragmatics», in *Journal of Pragmatics*, n. 43, pp. 2148-2159.

MAZZONE, Marco (2013a), *Automatic and controlled processes in pragmatics* in CAPONE Alessandro, LO PIPARO Franco, CARAPEZZA Marco (a cura), *Perspectives on Linguistic Pragmatics*, Springer, New York, pp. 443-467.

MAZZONE, Marco (2013b), «Attention to the speaker. The conscious assessment of utterance interpretations in working memory», in *Language & Communication*, n. 33, pp. 106-114.

MAZZONE, Marco (2014a), «Crossing the associative/inferential divide: ad hoc concepts and the inferential power of schemata», in *Review of Philosophy and Psychology*, n. 5, pp. 583-599.

MAZZONE, Marco (2014b), «A generative system for intentional action?», in *Topoi*, n. 33(1), pp. 77-85.

MAZZONE, Marco (2014c), «Language and action: A common intentional, generative, and inferential process», in *Reti Saperi Linguaggi. Italian Journal of Cognitive Sciences*, n. 1, pp. 165-178.

MAZZONE, Marco (2015a), *Razionalità pratica e linguaggio*, Corisco, Roma-Messina.

MAZZONE, Marco (2015b) «Constructing the context through goals and schemata: Topdown processes in comprehension and beyond», in *Frontiers in Psychology*, n. 6 (651), pp. 1-13.

MAZZONE, Marco (2015c) «Trivializing modularity. An associative-representational account of cognition», in *Epistemologia*, n. 38, pp. 201-215.

MAZZONE, Marco (2016), «What kind of associative and inferential processes? A response to Rubio-Fernández (2013)», in *International Review of Pragmatics*, n. 8, pp. 143-161.

MAZZONE, Marco (in press), *Why don't you tell it explicitly? Personal/subpersonal accounts of implicitness*, in CAP Piotr, DYNEL Marta (a cura), *Implicitness: From lexis to discourse*, Benjamins, Amsterdam.

MAZZONE, Marco, CAMPISI, Emanuela (2013), «Distributed intentionality: A model of intentional behaviour in humans», in *Philosophical Psychology*, n. 26(2), pp. 267-290.

MCCARTHY, John, HAYES, Patrick John (1969), «Some philosophical problems from the standpoint of artificial intelligence», in *Machine Intelligence*, n. 4, pp. 463-502.

PASTRA, Katerina, ALOIMONOS, Yiannis (2012), «The minimalist grammar of action», in *Philosophical Transactions of the Royal Society B*, n. 367, pp. 103-117.

RECANATI, François (2004), *Literal Meaning*, Cambridge University Press, Cambridge.

SHANAHAN, Murray, BAARS, Bernard (2005), «Applying global workspace theory to the frame problem», in *Cognition*, n. 98, pp. 157-176.

SIMON, Herbert A. (1992), «What is an explanation of behavior?», in *Psychological Science*, n. 3, pp. 150-161.

SPERBER, Dan (2005), *Modularity and relevance: How can a massively modular mind be flexible and context-sensitive?*, in CARRUTHERS Peter, LAURENCE Stephen, STICH Stephen (a cura), *The Innate Mind: Strucuture and Content*, Oxford University Press, Oxford, pp. 53-68.

SPERBER, Dan, WILSON, Deirdre (1996) «Fodor's problem and relevance theory: a reply to Chiappe and Kukla», in *Behavioral and Brain Sciences*, n. 19, pp. 530-532.

WILSON, Robert A. (2008), *What computations (still, still) can't do: Jerry Fodor on computation and modularity* in STAINTON Robert J., EZCURDIA Maite, VIGER, Christopher (a cura), *New essays in philosophy of language and mind*, Supplementary issue 30 of the *Canadian Journal of Philosophy*.

WILSON Timothy D. (2002), *Strangers to Ourselves. Discovering the Adaptive Unconsciuos*, The Belknap Press of Harvard University Press, Cambridge (Massachusetts) and London.

YEH Wenchi, BARSALOU Lawrence W. (2006), «The situated nature of concepts», in *American Journal of Psychology*, n. 119, pp. 349-384.