On the Necessity of Mixed Models: Dynamical Frustrations in the Mind

Diego Gabriel Krivochen

e-mail: diegokrivochen@hotmail.com

Abstract: In the present work we will present and analyze some basic processes at the local and global level in linguistic derivations that seem to go beyond the limits of Markovian or Turing-like computation, and require, in our opinion, a quantum processor. We will first present briefly the working hypothesis and then focus on the empirical domain. At the same time, we will argue that a model appealing to only one kind of computation (be it quantum or not) is necessarily insufficient, and thus both linear and non-linear formal models are to be invoked in order to pursue a fuller understanding of mental computations within a unified framework.

Keywords: dynamical frustration; Markovian models; quantum human computer; Turing-computation

1. Introduction: A brief history of "quantum mind" proposals

With respect to the scientific developments that led to the different versions of quantum theories of mind, Stapp (2009: 4) claims that

"This [quantum] model of the mind/brain system is no isolated theoretical development. It is the rational outcome of a historical process that has occupied most of this century, and that links a series of revolutions in psychology and physics."

Even if the historical antecedents that Stapp mentions go back as far as the 19th century, our brief review will start in more recent times. Already in the '40s it was clear to some that Statistical Mechanics and linear models could not account for the stability and (chaotic) ordering of natural systems (e.g., meteorology, which was the original field of Lorenz's studies), even within biology (Vitiello, 2001: 69; Schrödinger, 1944). The Cognitive Revolution of the '50s brought along a strong support for computational theories of the mind, and the formalism that outmatched the others was, by and large, Alan Turing's: to this day, there are Turing models of the mind (see, for instance, Watumull, 2012). However, the quantum revolution that had taken place in the early decades of the 20th century had influenced part of the field of cognitive studies, and the idea that quantum effects are not just oddities at the Planck scale (ultimately, an idea stemming from the EPR paradox and Einstein's research on relativity) began to grow and develop. In this scenario, cooperation between physicists and brain scientists (cognitivists and neurologists) started around 1960, with the possibility of conceiving the brain as a many-body system: there are subsystems and their repeated complex interactions create quantum correlations. This, incidentally, implied abandoning materialistic dualism as a philosophical stance: the clear-cut separation between brain and mind was not so clear-cut

anymore, despite the *reduction ad absurdum* arguments Gilbert Ryle had given in 1949 against unification frameworks (see Dennet, 1991 for discussion), partially based on the Cartesian idea that nature is to be divided in two non-related (and non-unifiable) parts: mind and matter¹. Local reductionism and determinism, characteristics of classical physics, were now questioned, particularly after the first observations of hypersensitivity to initial conditions (consider that the first 'chaotic' observations by Lorenz took place around 1963) and further developments in complex systems.

Going a step further from the many-body problem mentioned in the previous paragraph, inserting language (and the mind as a whole) in the natural world, as a physical system just like any other, allows us to dispense with the undesirable consequences of looking at it as a closed system (that is, a system which is insensitive to external factors): let us imagine that we have N (where N is a natural number) strings, using Chomsky's own terminology, and n automata (nevermind whether they are alive or not) making use of those strings. If interactions are binary (that is, only two automata are interacting at any given time T), the "cycle" it would take for a string to re-appear, that is, the total amount of possible states of the system of interactions of n "saying" N is defined by the expression 2^N (Bernárdez, 2001). Assuming only 50 instantiations of NP strings, that amounts to 2⁵⁰: pure statistics can do little to help in these situations. However, if we open the system (make it sensitive to external factors), a different kind of mathematics comes into play: chaotic mathematics (see Boccara, 2002 for discussion). As we have proposed in Krivochen (2013), let us assume language is in fact a non-linear, open system, hypersensitive to initial conditions and displaying a "many-body" like behavior. In these kind of systems, the wave function describing the state of the system holds a large amount of information, hardly manipulable by linear systems of the kind described by the Chomsky Hierarchy, for example (Chomsky, 1957; Lasnik, 2011; Lasnik & Uriagereka, 2012). Like many (other) nonlinear systems, in language the output values are not proportional to input values, which amounts to say that there is more in an LF representation than there is in a Numeration or Array: interface objects are made of lexical items plus structure, and structure is significative in itself (Hale & Keyser, 1997: 40).

The dependencies between the elements of the quantum system are non-local, that is, they can appear at a long-range, as first noticed by Einstein, Podolsky & Rosen (1935)². Crucially for a model of quantum linguistics (see Krivochen, 2011a, b, 2012, 2013), the relevant measurements over possible

¹ Interestingly, the first unification theories we have notice of come from the *empiricist* side, with philosophers like David Hume (see Chomsky, 2005b for some discussion).

² Einstein's description of quantum entanglement is almost as famous as the notion itself: "spooky action at a distance".

outcomes in a quantum system are binary³, we are always talking about pairs of measurements (Stapp, 2009: 5), which relates to an apparently essential property of phrase markers and constituents in general, at least at the interface level of semantic structure (C-I). Interestingly enough, the predicted interferences between experimental results on measurements are macroscopic phenomena, not Planckscale effects; and Vitiello reports neurophysiological evidence of long-distance neurological action, which cannot be explained by means of single-neuron models. Memory (information retrieving from the Long Term Memory LTM) seems to be an obvious example, and the evidence Phylyshyn (2007) presents in favor of distributed computation of Prepositional Phrases PP (in localist terms, figureground dynamics, see also Talmy, 2000) in the temporal and parietal lobes seems also an interesting path to take. The point made by Vitiello, echoing Freeman (2000) is that, even if it cannot be claimed that all neural connections and brain activity respond to quantum modeling, there are processes that just cannot be modeled in a traditional model. In recent years, not only studies in human neurophysiology but also AI (in a move that was somehow anticipated by Penrose, 1997) have attempted to generate a quantum theory of the mind (some, more inclined to so-called "consciousness"), maintaining the *computer analogy*. This, needless to say, required a deep revision of the fundamental assumptions of AI (unfortunately, to the best of our knowledge, there has been no such revision in computational linguistics, which remains strongly statistical and primarily descriptive) when the first advances in quantum computers saw the light, not too long ago. True, quantum mechanics is a statistical theory, but in a whole different sense: prior to observation / measurement, a particle's momentum (for example) is to be defined as a probability, not a certain datum. Moreover, the particle itself is not a little ball of non-divisible matter, but more likely a complex unit itself, product of the vibration frequency of 1-D strings at the Planck scale (1*10⁻³⁵ m), if (some version of) string theory is on the right track (see Greene, 1999 for discussion at an introductory level). This complexity in interactions gives rise to systems whose behavior cannot, foreseeably, be fully accounted for by classical (i.e., Newtonian) mechanics. The mind, it is argued by some (including us) is one of those systems. What is more, some mental systems (as we will argue, language among them), in the sense of symbolic structures generated by neurological processes display macro (i.e., observable) quantum properties of the kind mentioned earlier. This thesis, which is sometimes called "quantum human computer hypothesis" (QHC) is crucially independent of the narrower thesis that language itself is a chaotic system, which we have also put forth in previous works (Krivochen, 2013), in connection to the QHC. It is essential to point out that the two theses are independent, and it is possible to adhere to one without necessarily adhering to the other. For

_

³ This, in traditional quantum physics, derives from the so-called "wave-particle duality". We will see that this is not always the case, as we will work with elements that present more than two possible outcomes.

example, Uriagereka's (2012) CLASH model, based on the notion of *geometrical frustration*⁴ (see Binder, 2008 for details) is compatible with the second thesis (the chaos thesis), but major changes would have to be performed in the theoretical apparatus if the CLASH system is to be implemented in a quantum mind⁵. For the purposes of the present argumentation, and following the line of Krivochen (2011a, b, 2012a, b, c, 2013) we will simply characterize the "quantum human computer" as follows:

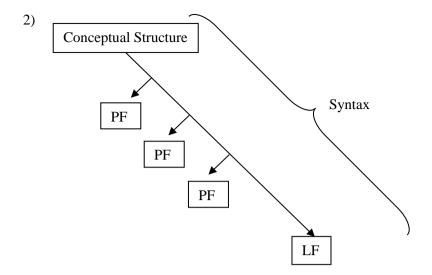
1)

- a. It is a computational system, which builds on the assumption that mental processes are derivational
- b. It builds on the assumption that derivations create representations that are evaluated by interpretative systems, which interface with the generator (GEN) algorithm
- c. It allows any object O of arbitrary complexity to comprise, before interpretation (i.e., transfer to the interpretative systems, whichever they are), n > 1 states at once. n collapses to one of the possible outcomes at the interpretative levels, not before.
- d. It is blind to the characteristics of the manipulated objects

The aforementioned assumptions are related to (even if in a non-necessary way) a proposal about the architecture of the cognitive system underlying language production and comprehension, and the mathematics necessary to model it. The architecture we assume is the following:

⁴ A geometrical frustration presents global and local tendencies which are mutually contrary. Binder (2008: 322) illustrates the situation with a Lorenz attractor, whereas in Uriagereka's model (and our own) global tendencies can be exemplified with semantic information (the CS-LF arrow in figure (2)) while local tendencies arise from a Multiple Spell Out model, and involve the materialization of locally determined chunks of structure (the arrows leading periodically to PF).

⁵ The adaptations that traditional models would have to undergo if the QHC hypothesis turns out correct is a fascinating matter in itself. Consider, for example, the following quotation from Stapp (2009: 18): "The fact that, for example, a certain pointer appears to any community of communicating observers to have swung only one way, or only the other way, not both ways at once, is understood in terms of the idea that the universe splits, at the macroscopic level, into various non-communicating branches" (emphasis in the original). It is obvious how the idea of non-communicating branches (i.e., not related by any dominance / sisterhood relation) impacts on phrase structure, particularly regarding the displacement property of human language. See Krivochen (2013) for discussion, but the matter is far from being solved.



In our terms, a derivation does not start with a Numeration (a set of elements with numerical subindexes indicating how many times they will be used in a derivation, see Chomsky, 1995), but with a pre-linguistic purely conceptual structure, in the line of Fodor (1975) and, more recently, Jackendoff (2002), Culicover & Jackendoff (2005), Uriagereka (2008), and the sense in which D-Structure is understood in Uriagereka's (2012) CLASH model. That structure is syntactic in a wide sense, as concepts are structured (taking "syntactic" not in the narrow sense of "linguistically structured" but in a strict sense of "structured". This conceptual structure, shaped by the speaker's intention to convey a certain propositional meaning through linguistic means, is what, in our proposal, drives *Select*, the selection of a subset of LEX, in turn a set of linguistic *types*, to be instantiated as *tokens* in the "syntax" (actually, not a component but a workspace, in the sense of Baddeley, 2003) driven by the need to *minimize entropy* as the derivation unfolds. The assumption we make at this respect is the following:

3) Minimal Selection:

Select the minimal amount of types that can instantiate a conceptual structure CS into a linguistic structure LS losing as few information as possible.

_

⁶ Cf. Culicover & Jackendoff (2005: 20 fn. 8): "Algebraic combinatorial systems are commonly said to 'have a syntax'. In this sense, music has a syntax, computer languages have a syntax, phonology has a syntax, and so does Conceptual Structure. However, within linguistics, 'syntax' is also used to denote the organization of sentences in terms of categories such as NP, VP, and the like. These categories are not present in any of the above combinatorial systems, so they are not 'syntax' in this narrower sense." In this paper, and in general within our theory, "syntax" is used in the wider sense, for two main reasons: to begin with, there is no compelling evidence that the "syntactic mechanisms" vary from one system to another (except insofar as the units affect the algorithm, in case that actually happens); and also, an adequately wide formalization of syntactic mechanisms could reveal deep facts about the structure of more than a single system. Admittedly, this requires interdisciplinary co-working and terminology unification, which are unfortunately not the norm now.

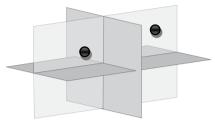
The intuition behind this assumption is clear: we want to linguistically instantiate a CS in the most economical way possible⁷, ceteris paribus. Given the fact that the CS includes not only rough propositional content but also added information (what most linguists would put under the "pragmatics" label: inferences, and other extra-propositional which is, nonetheless, built upon the clues syntactic structure provides the semantic component with), the reference set for each potential derivation is unary: there is one and only one candidate which can express CS in an optimal way. Assuming the existence of (some form of) a Lexicon for human language, Select, then, builds an array of lexical types from that Lexicon. Then, units are blindly manipulated in the workspace via concatenation:

4) Concatenation defines a chain of coordinates in n-dimensional generative workspaces W of the form $\{(x, y, z...n) \subset W_X, ...(x, y, z...n) \subset W_Y, ...(x, y, z...n) \subset W_n\}$.

Simplifying the matter almost excessively for the sake of clarity, take "dimensions" to mean the number of coordinates necessary to define the position of a point. Thus, each set of coordinates depends on the number of dimensions in the relevant generative workspace, such that an element is to be defined by all of its coordinates in W (that is to say, there are no "superfluous" coordinates in a dimensional specification). We assume only one condition for any X and any Y to enter the concatenation relation: they must share what we have called "ontological format": ontological format refers to the *nature* of the entities involved. For example, Merge can apply ("ergatively", as nobody / nothing "applies Merge" agentively) to an *n* number of *roots* because they are all linguistic instantiations of generic concepts (Krivochen, 2011a: 10; Boeckx, 2010). With ontological format we want to acknowledge the fact that a root and a generic concept cannot merge, for example. It is particularly useful if we want to explain in simple terms why Merge cannot apply cross-modularly: a root and a phoneme do not share ontological format (they have different nature, one conceptual, the other phonological), therefore, the system blocks such an operation from square one.

Given this scenario, let us see how an XP would be formed, say, a DP (assuming the simplest possible structure: [D, $\sqrt{}$]):

⁷ In more technical terms, Selection must *reduce entropy*. If the theory of Merge we have developed in past works is correct, the generative algorithm, driven by interface requirements, should also be "counter-entropic" (see also Uriagereka, 2011). The possibility is currently under research.



Workspace 1 (W1)

Both D and $\sqrt{}$ having the same ontological format, Concatenate can (and thus must) apply in the following form:

6) Concatenate (D,
$$\sqrt{}$$
) = {(x, y, z) \subset W1, (x', y', z') \subset W1)}

The coordinates of the result of the operation (a DP, or {D}, construction) are defined as the Cartesian product of the (in this case) two sets of coordinates of the elements involved in the merger. In the more familiar tree form, the result would be represented as (7):



A note is in order here, particularly taking into account the discussion in sections below: the newly formed syntactic object, even if irrelevant for the generative algorithm as such, must be identified as a unit for the purposes of further computations, what is customarily referred to as a "label". In past works (mainly, Krivochen, 2011a) we have argued against the existence of labeling in the syntactic workspace, primarily given their null pertinence to the derivation being the algorithm both free and blind. This means that, if existent at all, labels are only relevant at the LF interface (since it is very difficult to argue how labels could be of any interest or relevance for PF purposes). Instead of providing a stipulative labeling algorithm, based on alleged UG principles (Chomsky, 2005a; Gallego, 2007), we claim that the label of an object is nothing more than a "summary" of its semantic properties, which, just as categories, or Case; is recognized at the interface as the result of a configuration. Gallego (2007: 75) claims that [in a Merge (α , β) situation] "we cannot know whether β is a LI or an XP [...] without labels". Our objections to this position are simple: a) at the syntactic workspace, it is not necessary to know it, because the algorithm is blind; and b) the label is the reading of a configuration, not the other way around. Having D and $\sqrt{\ }$, should C-I "label" as $\sqrt{\ }$, there would be a crash, since the root is too semantically underspecified to be used to refer (either to a sortal or an eventive entity). The only way out is to recognize the whole construction as a D, a sortal

entity. In this sense, we dispense with labeling algorithms like those summarized in Gallego (2007) and including Chomsky, Boeckx, and Hornstein; and propose a theory that is even simpler than the "label-free" alternative of Collins (2002), as we do not need the notion of *locus* (which ultimately amounts to selection). In any case, the labeling discussion is well outside the study of dependencies in the generative workspace.

We would like, at this point, to make our architecture crystal-clear. We base our theory, like Culicover & Jackendoff (2005); Uriagereka (2012), among others, on a pre-linguistic, syntactically built conceptual structure, which has to be instantiated via language, considering requirements and limitations from both phonology and semantics. However, complementarily to Uriagereka (2012), we focus on the semantic side of the story, and explicitly state the preeminence of semantics over phonology for conservation (i.e., anti-entropic) purposes. As we will see, most of the problems we find hard to solve from a Turing-computer perspective arise when one goes beyond inferring syntax from phonology (as Kayne, 1994; Moro, 2000, and much subsequent work do). We adhere to Uriagereka's (1999) Multiple Spell-Out model, which implies that access to the phonological interface (or, in our terms, access from the phonological interface to the syntactic workspace) is performed multiple times within a derivation, thus basing the computation on the notion of local cycle, and extend it also to the semantic interface. The difference with Chomsky's (1998, 2005a) phase-system is that Urigereka's proposal, and our own, are based on interface requirements (in Uriagereka's case, the impossibility of linearizing determined phrase markers), which, if the interfaces are independent, means that PF phases and LF phases need not coincide (contra Chomsky, 2005a, even though references to the matter in Chomsky's work are too vague to constitute a stance). The derivational dynamics we will assume hitherto (summarizing points and discussion made in previous works, see Krivochen, 2011a, b, 2012a, b, 2013) is as follows:

8) Concatenate (α, β) = {α, β}
 Analyze_{IL} {α, β} [is {α, β} fully interpretable by an interface level IL?]
 (Transfer {α, β]} to IL if Analyze_{IL} results in convergence at IL)

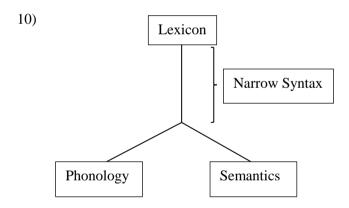
At each transfer point there is a tension: unlimited syntactic resources (after all, *concatenation* can apply unboundedly) and limited materialization possibilities (given by the array of phonological exponents available in a given language L). This, if we follow Binder (2008) and Uriagereka (2012), gives rise to a *frustration*, on which the whole system is built. On a similar line, we will assume a strong optimalization thesis, to be (informally) formulated as follows:

9) Every externalized linguistic object E is the optimal resolution of the geometrical frustration involving the global infinitude of syntax and the local (un)availability of phonological exponents in L.

Our goal in this paper will be to give evidence in favor of the thesis that some processes (at least) cannot be Turing-computable or even modeled by a simple, linear L-grammar. We will focus on two such cases (while mentioning others in the conclusion, for reasons of space): categorization, and case.

2. Remarks on Categorization

Chomsky's (1970) *Remarks on Categorization* (RC) have the strange merit of being considered the foundational stone for two opposite conceptions about syntactic categories: lexicalism and distributed morphology. On the one hand, we have a theory that assigns the Lexicon generative power to different extents, from the GB-influenced L-Syntax of Hale & Keyser (1993) to the highly developed non-transformational model put forth by Ackerman et. al. (2011), the so-called "implicative morphology". In any case, the basic thesis of lexicalism is that syntactic mechanisms do not make reference to word-internal processes, nor can they manipulate smaller-than-words constituents, be them morphemes or roots. In one form or another, lexicalism assumes the Y-model, depicted in (10):



The "syntax" lexicalism often refers to is the so-called "narrow syntax" (Hauser, Chomsky & Fitch, 2002), which builds symbolic representations from lexical items, at that point opaque to external influence. Elements enter a derivation as sets of features (an assumption shared by Minimalism and non-transformational models, like HPSG or LFG), including semantic and phonological features, as well as, in some cases (e.g., Green, 2011) syntactic specifications regarding subcategorization frames (quite like GB lexical entries, but considerably richer). Two tendencies can be distinguished, broadly speaking: for some (see Williams & Di Sciullo, 1987; Lasnik, 1999; Solá, 1996; Green, 2011), lexical items enter the derivation fully inflected, perhaps with some exceptions (verbs [be] and [have], in Lasnik's proposal). This thesis is sometimes called "Strong Lexicalist Thesis", and claims that both inflection and derivation belongs to a module which is separate from the syntax, ruled by different

principles. For others, including Aronoff (1976) and Chomsky (1998), Case and Tense inflection are processes that take place within the Narrow Syntax (NS), in the case of the latter via feature valuation (see De Belder, 2011: 22, ff. for comparison and discussion). Chomsky considers that features enter a syntactic derivation either valued or unvalued, depending on the category they are part of. Thus, Person/Number are inherently valued in N and Pronouns, whereas they are unvalued in V. Since unvalued features cannot be interpreted by the interfaces PF and LF (Chomsky, 1999), unvalued features are assigned a value during the course of the derivation and then, according to some proposals (e.g., Kitahara, 1997), erased (but see Epstein & Seely, 2002 for powerful arguments against the notion of erasure). Needless to say, Chomsky's system requires categories to be fixed in the Lexicon, a stipulation that comes concomitant to that determining which features are valued in which category. However, this is, to the best of our knowledge, not a way to solve a problem, but merely to wipe it under the rug. Problematization came from lexical decomposition perspectives, Distributed Morphology (Halle & Marantz, 1993), and Exo-Skeletal Models (Borer, 2005, 2009). The common denominator to these approaches is that categories arise as the result of interactions within the syntactic workspace (see De Belder, 2011 for discussion). The issue, complicated though it might seem, can be exemplified very easily. Consider (11):

11) $\sqrt{\text{water}}$

We have used an English word to stand for the root content, but it is worth noting that roots are language-neutral, that is, the set of roots is most likely universal. Now consider the two following contexts:

- 12) a. John watered the plants
 - b. John drank a glass of water

We have two options: either we posit that the Lexicon has two fully-fledged (i.e., already categorized and with some fixed features) entries, \mathbf{water}_V and \mathbf{water}_N , or we assume that there is a root $\sqrt{\mathbf{water}}$ that somehow acquires category in a specific context⁸. Lexicalism assumes the first option, we assume the second on empirical and theoretical grounds. One of the strongest arguments in favor of "post-syntactic categorization" is the existence of not only categorial, but also argumental alternances. For example:

- 13) a. John broke the glass
 - b. The glass broke

_

⁸ Examples analogous to (12) are easily found on Hale & Keyser (1993, 2002), Mateu Fontanals (2002), and related work on lexical decomposition and argumental alternances.

And so on. In a strong lexicalist model, we would have not only N and V diacritics within the lexicon, but also some notation to differentiate [break $_{ERG}$] from [break $_{CAUS}$]. That notation would go directly against any Occam-related desideratum, since entities (in this case, lexical entries) would be multiplied beyond necessity (if we can come up with a more economical theory). Before getting fully into the topic, let us make explicit some assumptions we will draw upon during our inquiry:

- 1) Categories, phases and other units are not primitives of the syntactic theory, but arise as a result of the interaction of a free Merge system with interface conditions: the dynamics of the derivation and the legibility conditions of certain interpretative mental faculties or any other computational module. (see Krivochen, 2012; De Belder, 2011, Boeckx, 2010; also work in Distributed Morphology like Marantz, 1997 and Fábregas, 2005 and Exo Skeletal Models, see Borer, 2005, 2009 among others).
- 2) There is no distinction between "lexical derivations" and "syntactic derivations", and this goes beyond positing a single generative mechanism: there are just derivations, regardless the nature of the elements that are manipulated, since the generative operation is blind. This means that there is no pre-syntactic generative lexicon (Cf. Pustejovsky, 1995; Hale & Keyser, 1993) and no constraints on Merge (Cf. Chomsky, 2005a and his "Edge Feature" as a *sine qua non* condition for Merge to apply; also Pesetsky & Torrego's 2007 *vehicle requirement on Merge*; Wurmbrand's, 2013 *Merge Condition*, among many others). For the historical basis of this claim, see Halle & Marantz, 1993, and subsequent work in Distributed Morphology.

Our reasoning goes as follows: if a root $\sqrt{}$ can be externalized as X, Y...n, then it must bear the potentiality to have those functions. In other words, if a root can surface as either an N, an A or a V, then it must have the potential to be an N, an A and a V. What is more, prior to a specific derivation, in isolation, the root's status can be described, following a very well-known convention in physics first formalized by Erwin Schrödinger, as the addition of the possible outcomes, configuring a "wave function" instead of locating the root within the cognitive workspace in terms of classical coordinates (see, e.g., Langacker, 2007; Talmy, 2000, 2007). The structure of the lexicon, thus, is to be deeply revisited, insofar as so-called "lexical categories" (or "conceptual categories", in a more Relevance-oriented framework, see Escandell Vidal & Leonetti, 2000 for discussion) can be seen as roots in their " ψ -state" (i.e., comprising all possible outcomes, following Schrödinger, 1935, Section 5). This simplifies the lexicon enormously, as, for instance, [shelf_N] and [shelve_V] are grouped under a single entry, [$\sqrt{}$ shelf]. But how do roots get "categorized", then? We find two possibilities:

a) Via Merge with specific category-defining functional heads, like *v*, *n*, *a*; etc. (Marantz, 1997; Fábregas, 2005; Panagiotidis, 2010).

b) Via interface reading of a *local* dependency between a root and a functional head not specifically devised for categorization purposes.

The difference is great in both theoretical and empirical domains: the first approach needs "categorizers", functional heads whose only contribution to LF is to provide category to the roots they have scope over. However, this does not solve the problem, it is simply a stipulation, as sometimes those alleged "categorizers" have no impact over PF (that is, they are not realized as morphemes) and sometimes they are, depending not only on the language (e.g., English is much more inclined to conversion than Spanish) but also on the relevant root, a difference that is left unexplained in the literature about categorization we know of. It is also quite an anti-minimalist answer, since it assumes a functional head *per* "part of speech" (see Fábregas, 2005: 32). In the second proposal, we have a very narrow set of semantically relevant functional elements, which in other works we have made explicit as *v* (comprising *causativity*), T (comprising *time*), P (comprising *location*), D (comprising *sortal referentiality*) and C (comprising *illocutionary force*). What is more, if the syntactic component is as underspecified and blind as we have characterized it, then there is no place for "categories" there: they must arise at the LF interface, *after* transfer. We claim that a category is the result of a local relation between a root and a distributionally specified functional head. But, which are the correct correlations? Let us take a quote from Aristotle's Poetics:

"A Noun is a composite significant sound, not marking time [...] A Verb is a composite significant sound, marking time, in which, as in the noun, no part is in itself significant. For 'man', or 'white' does not express the idea of 'when'; but 'he walks', or 'he has walked' does connote time, present or past". (Aristotle, Poetics XX, 8-9)

Needless to say, there are more recent references to the matter, but no doubt less clear and stained with some theoretical framework or the other. This fragment presents a fact, which in more contemporary terms could be rephrased as "there is no T node within DPs". This is already something, since if T is absent from DPs, it cannot be T that categorizes a root as N. On the other hand, and in parallel, there is no D within an eventive structure. Summarizing the discussion made in Krivochen (2012: 90, ff.), T is distributionally specified enough to generate an eventive reading, and D is distributionally specified enough to generate a sortal reading. So far, we have derived two tyoes of entities, sortal (N) and eventive (V), but what about properties of those entities (Adj. and Adv.)? In this respect, we follow the localist theory of Talmy (2000) (also adopted in Jackendoff, 1987) and the lexical decomposition perspective explained in length in Mateu Fontanals (2002) and Hale & Keyser (2002), among others. From the combination of these perspectives there follows the conclusion that both Adverbs and Adjectives are abstract locations in unaccusative conceptual structures, therefore prepositional in nature. Let us give an example:

14) Mary is beautiful

[
$$_{V}$$
 BE [$_{P}$ Mary [[WITH] $\sqrt{\text{beauty}}$]]]

15) Berlin is far away

The prepositional node, which can adopt two values (central – terminal coincidence), relates two entities in a *figure-ground* manner (Hale & Keyser, 2002: 218). Properties of entities (be them sortal or eventive) are *grounds*, syntactically located as complement to the P head (Hale & Kayser, 2002: 47, ff.). Being that P phonologically defective, it triggers *conflation* of its sister, which is sometimes spelled out as an affix (e.g., beautiful = with+beauty).

Let us now express what we have discussed above in a more schematic form:

16) A lexical item LI is a structure $\{X...\alpha...\sqrt\}$ \in W_X , where X is a distributionally specified functional category (Determiner, Tense, Preposition), α is an n number of non-intervenient nodes for category recognition purposes at the semantic interface, and \sqrt is a *root*.

And the correlations result in the following distributional patterns:

17) a. N =
$$[D...\alpha...\sqrt{}]$$

b. V = $[T...\alpha...\sqrt{}]$
c. A / Adv = $[P...\alpha...\sqrt{}]$

where α is an n number of non-intervenient nodes for Minimality purposes, because they are not distributionally specified enough. Let us see some cases: v is, in our opinion, not specified enough to generate a categorial interpretation at the semantic interface (thus collapsing the root's ψ -state), because it can appear in both sortal and eventive contexts, if the sortal entity is a derived nominal. For example:

18) a. The enemies destroyed the city

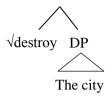
b. The enemies' destruction of the city

Let us analyze the derivation step by step.

19) a. We start with a DP [the city], which is merged with a node [√destroy], underspecified as regards category. Since our generator function is blind and free, there is no featural

⁹ Following Escandell & Leonetti (2000), we assume that functional categories are procedural insofar as they provide the semantic interface with instructions as to how to interpret the relation between entities over which they have scope.

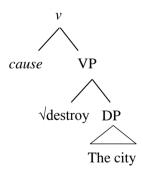
requirement whatsoever to trigger Merge (Cf. Wurmbrand, 2013; Pesetsky & Torrego, 2007, among others), therefore the merger of a root and a DP is not banned in principle.



b. So far, we have a sortal entity [the city] and a root generically denoting an event. The label, for C-I purposes, is then VP, as the "projection" has been closed since the next derivational step will introduce a different kind of information ¹⁰ (but see Krivochen, 2011a, 2012 for discussion about the possibility of having a different labeling system, dispensing with bar-notation).



c. Next, we introduce another semantically interpretable element, the primitive *cause* (see Mateu Fontanals, 2005 for discussion). The construction is thus read by C-I as a *caused transitive event*.

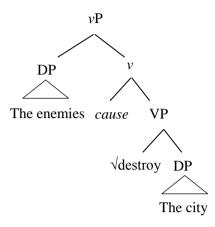


d. The primitive *cause* requires the introduction of an actant in the construal: an *initiator* (independently of the presence of an object, consider for example unergative verbs). A further structural position is licensed, where a DP is merged and interpreted thematically as the agent/initiator of the event over which the primitive *cause* has scope. The causative projection

14

¹⁰ Admittedly, this step requires some look-ahead, which is a problem for real-time labeling under traditional assumptions. For reasons of space, we have not discussed labeling in a system of invasive interfaces, as we do in Krivochen (2011a, 2012), but we refer the reader to those works for details.

is then closed, since there is no more information of the same nature (i.e., eventive / causative) to add to the construal.



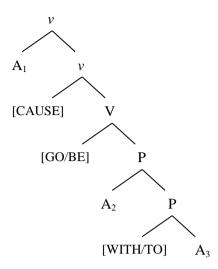
e. So far, nothing has been said about category recognition, and this is because, up to this point, there is no certainty about the distribution of the construction. For all we know, it could be either "the enemies destroyed the city" or "the enemies' destruction of the city", since those constructions have both (semantically speaking, and for all that matters) the same underlying construal: a caused transitive event. Neither V nor v are distributionally specified enough to generate a categorial interpretation at the semantic interface, which means that, up to this point, the whole vP is in a ψ -state as far as category is concerned. This is important because it means, should it be true, that the syntactic workspace can host a structure of arbitrary complexity in its ψ -state, comprising all possible outcomes, and for as long as necessary. If transfer is nothing more than the interfaces taking from the workspace the minimal units they can read (and not the syntax sending information to the interfaces, as in Chomsky's 1998, et. seq. proposals), then, in principle, there is no limit to the amount of non-Markovian / non-Turing computable structure that can be kept active. Of course, there are issues of memory, but that is quite another problem, having little to do with computational capacity (consider, for example, that Turing machines are claimed to have unlimited memory, see Uriagereka, 2012: 230-231; yet they are clearly unable to process non-linear dependencies, as we would find in a Lorenz attractor and, perhaps, even in human language, see Krivochen, 2013 for discussion). If there is a geometrical frustration deep inside language design, then we have to add a level to the Chomsky hierarchy, to include non-classical computation, among which we count quantum computation.

To summarize, until a distributionally specified node is inserted in the structure, be it D, T or P, the state of the symbolic object in hand is to be described as the "sum" of all possible outcomes, comprising many possible states at once as potentialities. This, we argue, is only modelable by means of quantum computations.

Going beyond the word-level, the Case-Theta system also offers a good example of a "many possible outcomes" situation. The case for Case we have made in previous works applies here as well, so we will summarize our arguments and refer the reader to those works for more discussion and examples. To begin with: what is Case? Does it have any syntactic relevance? Our answer to these questions are somehow one and the same: Case is, just as category, an interface reading of a syntactic configuration. Just like category, also, we need particular procedural nodes that convey the relevant instructions for C-I to read and interpret. That, as we have said, is one cycle. The other, morpho-phonological cycle, is where, as many have claimed (within and outside Chomskyan orthodoxy), inter-linguistic variation lies¹¹. The morphological realization of Case as a morpheme, despite some inter-linguistic regularities (e.g., the <m> is associated to Accusative in Latin, English, and German plural), is an epiphenomenon as far as syntactic-semantic processes are concerned. Which are the relevant processes, then? At this point, we would like to introduce an interesting parallel between the Case/Thematic and categorial systems we have explored in past works (mainly, Krivochen, 2011a, 2012a; Krivochen & Luder, 2012): they are both interface readings of configurations of the kind $[X...\alpha...Y]$, where X is a procedural node, α is an *n* number of non-intervenient nodes and Y is an object of arbitrary complexity, more specifically an entity, either sortal of eventive. Case, as it is obvious, affects only sortal entities, which can, in very broad terms, either affect or be affected. This semantic distinction leads to the binary Case systems, *nominative-accusative* and *ergative-absolutive*. Those labels, however, refer to the morpho-phonological cycle, and notions of markedness (e.g., which is the unmarked Case in L?) which have no place in a semantic approach. Consider now the following scenario, partly depicted above: there are two event-related nodes that take arguments (following Hale & Keyser, 2002; Mateu Fontanals, 2002, 2005), namely, v (the causative node requires an *initiator*, realized categorially by means of a sortal entity) and P (the locative node relates a figure and a ground, both sortal entities). The V node is a transitional node, which conveys Aktionsart-related information (that is, if the event is dynamic or stative), but takes no arguments. This leaves us with the following structure:

¹¹ Above, we have referred to a global semantic tension and local phonological tensions. Consider, then, semantics as a macro-cycle and phonology as micro-cycles, with opposing tendencies. There, a geometrical frustration arises.

20)



We have three structural positions available for arguments, all, as we have said, associated with a specific semantic interpretation. At this respect, DeLancey (2001, Lecture 3) says:

"(...) suppose we could demonstrate that there are, say, exactly x universal semantic roles which can occur as core arguments in a clause in human language. The most obvious language design would have x case markers, one for each underlying role; every argument would simply be marked for its semantic role, which could then be read directly off the surface morphosyntax (...)" [our highlighting]

While it has already been pointed out that "surfice morphosyntax" has little to do with the problem of Case (Spanish, for instance, only marks ACC and DAT Case on pronouns and clitics, but abstract Case, in the sense of Vergnaud, 1977), the intimate relation between Case- and Theta-positions is a strong point in De Lancey's presentation, and in ours (see also Krivochen & Luder, 2012 for discussion). From this paragraph, we conclude that, should there be at most three argumental positions, there are only three possible Case-Theta positions at most, in case we are dealing with a ditransitive structure. Inter-linguistic variation regarding the availability of Vocabulary Items to be inserted in terminal nodes and materialize Case (in a separationist framework, see Halle & Marantz, 1993 for the first developments of the notion of "late insertion") seems to go against the eliminative proposal of De Lancey, quite minimalist in spirit (way more than, for instance, Pesetsky & Torrego's 2004, where stipulations over feature valuation complicate the scenario beyond both necessity and desirability). Consider the Chomskyan proposal: if Case is an unvalued/uninterpretable feature, and those are valued (and thus made interpretable) via probe-goal relations with functional categories, a system like Sanskrit's would require eight distinct functional categories, one per "surface morphosyntax" expression of Case. Same happens with Latin's "6 Cases", or Ancient Greek's 5. We have argued in past works that there are only three fundamental Cases, structured as spheres, with a

prototype-periphery semantic dynamics (Krivochen, 2011a, 2012a: Chapter 2; Krivochen & Luder, 2012). In this framework, the three spheres are NOM, ACC and DAT, more accurately dubbed Initiator Case, Theme Case, and Location Case. As the reader may have noticed, we keep the "semantic preeminence" thesis, making reference to the semantic contribution of an element X in a position P to the LF rather than to morpho-phonological characteristics. With respect to the *spheres*, it is clear that the prototypical NOM occurrence is as an Initiator, structurally, Spec-vP, and there is nothing else you can do with it: NOM is, in all systems, the most distributionally constrained Case. ACC, on the other hand, may appear as either object in a transitive structure, or subject in an *accusativus cum infinitivus* clause, thus overlapping with what we would expect from NOM. The ACC sphere also includes those instantiations of elements that are semantically Themes moving towards a Location but displaying different morphological marks (e.g., Instrumental Case). DAT sphere includes all locative-like Cases, that is, all Cases in which there is a locative relation established between two entities, be it movement (*unde, quo, qua*) or possession. Thus, DAT sphere semantically includes morphological Locative, Genitive, and Ablative (Krivochen, 2012a: 79, ff.).

Going back to the diagram in (20), if there is a P involved, then there is locative meaning in the construal, and the complement of that P is the *ground* in the localist dynamics (Talmy, 2000; Anderson, 1977, among others). That *ground* corresponds to a Location, either literal (a place) or metaphorical (a property). Therefore, it is quite safe to assume that a local relation with P is the condition for the DAT sphere to be interpreted at the semantic interface in a particular DP. The *figure*, that is, the Theme that moves towards a Location, varies between NOM sphere and ACC sphere depending on whether it is an affected object or not: if we are dealing with a *caused* construal, then the *figure* in local relation with v will license ACC, if the construal is uncaused (e.g., unaccusative), the next functional element is T, licensing NOM. The final reflection is quite the same as in the previous section: if a DP can adopt any of the three spheres as a final state, it must bear the potentiality in isolation. Therefore, prior to the merger of v, P, or T, the Case-Theta status of a DP is, in the sense specified above, quantum. Summarizing:

21) Nominative: read off from a {Time, {D}} local relation, and thematically interpreted as Agent / Force

<u>Accusative</u>: *read off* from a {Cause, {D}} local relation, and *interpreted thematically* as Theme, the object (Figure) located in / moving towards, etc. a Ground.

<u>Dative</u>: read off from a {P, {D}} local relation, and interpreted thematically as Location, the Ground in Talmy's terms.

The inner complexity of the relevant quantum object (say, a DP) is nothing for the "syntax" to worry about, if by "syntax" we just mean a generative, multipurpose workspace generated *ad hoc* via

(according to D'Espósito, 2007) the activation of the pre-frontal neocortex and other relevant areas of the brain (e.g., temporal and parietal lobes, in the case of localist structures, see Pylyshyn, 2007). However, it would be too strong a hypothesis to claim that all mental processes share the quantum nature of language, which is partly due to the fact that there are two kinds of systems involved: generative and interpretative. Generative systems, being free and blind, can maintain and manipulate quantum objects, whereas *transfer* to interpretative systems collapses those objects to one of the possible outcomes. Not all subsystems in the mind work this way, and not even every linguistic computation is quantum, however. In the next section we will explore this possibility, which will ultimately lead us to a mixed model in which different processes involve different kinds of computations, either Markovian or non-Markovian; linear or quantum.

3. A Mixed Mind

The preceding discussion touches on an interesting point, namely, there are "macro" processes in which a quantum approach seems unavoidable. The scale of the modeling is essential for any argumentation regarding quantum computation in the human mind, since otherwise it is exposed to Litt et. al.'s (2006: 1-2) criticism regarding relevance of quantum considerations for mental phenomena:

"We argue, however, that explaining brain function by appeal to quantum mechanics is akin to explaining bird flight by appeal to atomic bonding characteristics. The structures of all bird wings do involve atomic bonding properties that are correlated with the kinds of materials in bird wings: most wing feathers are made of keratin, which has specific bonding properties. Nevertheless, everything we might want to explain about wing function can be stated independently of this atomic structure. Geometry, stiffness, and strength are much more relevant to the explanatory target of flight, even though atomic bonding properties may give rise to specific geometric and tensile properties. Explaining how birds fly simply does not require specifying how atoms bond in feathers."

If any, the contribution we would like to make here and in our past works (Krivochen, 2011a, b, 2012a, b) is that quantum phenomena can be found *beyond the Planck scale*, in mental computations¹². With categorization and Case-Theta interpretation we have provided an example that, even though accounted for with current theories (with different degrees of descriptive and explanatory adequacy), serves our purpose insofar as our explanation is, we believe, theoretically simpler and at the same time empirically robust, as it allows for coinage of neologisms and conversion just as long as the result is C-I interpretable.

We have reached a point in which we can say "there are at least some processes whose explanation requires an element to be described as a wave function". However, there is a missing part of the picture: are there all processes quantum within the mind? Our provisional answer, pending much

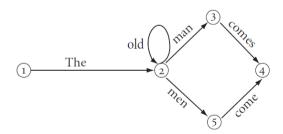
_

¹² Quantum effects beyond the Planck scale within physics have been identified, as we have said, since EPR's seminal work.

research, is no. Beyond Litt et. al.'s case against quantum models based on "consciousness" and "mathematical thinking" (which we will not discuss here, at least not directly), we will analyze linguistic dynamics that do not seem to require quantum explanations. This is only natural if we consider a fundamental geometrical frustration on the basis of generation-interpretation dynamics: global and local tendencies go in opposite directions (Binder, 2008: 322; Uriagereka, 2012). If there are quantum phenomena in language, then there *must* be Markovian (or other kind of classically computable) phenomena in the same system, thus configuring the opposing tendency. The claim that "quantum properties are irrelevant to explaining brain functions" (Litt et. al., 2006: 2) is, in our opinion, too strong. At this point, it cannot be denied from square one that there might be quantum phenomena in the mind, particularly taking into consideration the evidence proposed by the authors we mentioned in the first part of the present work. What is more, provided the thesis of geometrical frustration is on the right track (a matter still to be solved), there would be a strong architectural argument in favor of both quantum computation and traditional computation in the mind, without the need to dismiss any possibilities of non-linear computation. It is not clear, for instance, how Litt et. al. would deal with phenomena like categorization or multiple-candidate filtering in an OT-like architecture if not allowing the processor to perform multiple tasks at once and maintaining elements in a ψ -state until transferred.

In this section we will discuss the opposite tendency, exemplified by means of Markovian structures. Markovian models were claimed to be insufficient to account for *all* grammatical processes in Chomsky (1957), but this does not mean that *parts* of the grammar (e.g., specific constructions, if one adopts a Construction Grammar approach) cannot be Markovian. There are apparently two clear cases documented in recent literature (but drawing on old theories, going back to the '40s): *iteration* and *adjunction*. The case for iteration is simple: pure repetition (without semantic or syntactic scope involved) is better described as Markovian loops than by using phrase structure diagrams. For instance (see Uriagereka, 2008, Chapter 6; Lasnik, 2011: 355, ff. and Lasnik & Uriagereka, 2012):

22) The old, old,...man/men come/s



However, a Markovian syntax for such instances may not capture the semantic properties of some specific iterative constructions. Take, for example:

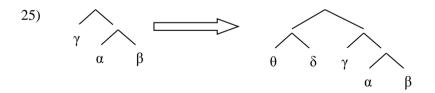
23) María es una mujer, mujer (Spanish)

'Mary is a woman, woman'

The meaning of this construction is not merely derived from the iteration, but, idiomatically, it means something like "Mary is very femenine". The power of Markovian explanations for iteration rests, partly, on whether idiomaticity is to be regarded a semantic or a syntactic effect. In our opinion, since semantics is syntactically structured, there is no choice but a mixed explanation, which takes into account the syntax-semantics interface (as partially done in Uriagereka, 2008, Chapter 6).

Anticipating discussion from Krivochen (in preparation), in turn heavily based on Uriagereka (2005, 2012), Markovian structures also seem to be relevant for Spell-Out purposes. In Uriagereka's (2012: 53) terms Finite State grammars find their limits in monotonic Merge, which is the application of the generative function in a successive way involving always a terminal node:

We see that the third step involves the inclusion of a terminal (i.e., non-branching node) γ which is merged with a non-terminal, $\{\alpha, \beta\}$, and the same happens in the fourth step, where φ is merged to a non-terminal $\{\gamma, \{\alpha, \beta\}\}$. The mechanism represented in (24) exemplifies this kind of application of the generative algorithm, which Uriagereka calls *monotonic*. *Non-monotonic* merge involves *two non-terminals*, as in (25):



In (25) we see that the second step involves the merger of two non-terminals, giving rise to a complex object. Each non-terminal, in turn, has been assembled by *monotonic Merge* in a separate workspace, and the unification takes place in a third workspace (in our proposal) or at the interfaces, after Spell-Out (in Uriagereka's). Relevantly, it seems that phonology works with Markovian dependencies (see Isardi & Raimy, in press), which means that both monotonic and non-monotonic structures (whose mathematical properties will not be discussed here) are to be "Markovized" via Spell-Out to be readable by S-M. This means that Spell-Out is nothing but dynamic markovization of non-Markovian material (e.g., complex lexical structures like path-of-motion and resultative predicates) or re-Markovization of elements that enter a workspace already in a "finite state grammar" format (e.g.,

adjuncts, according to Uriagereka, 2005), having been formed via monotonic merge in a separate workspace. This means that, according to the theory so far sketched, there are two kinds of Markovian objects in a linguistic derivation:

- a) Those derived by monotonic merge in a single workspace W_X
- b) Those derived in W_Y (where $X \neq Y$) and non-monotonically merged to Markovian objects derived in W_X

Taking into account Isardi & Raimy (in press), they must undergo a further process of Markovization, Spell-Out. They distinguish three "modules" of linearization, with different characteristics (Isardi & Raimy, in press: 3):

26) Module Characteristics

Narrow syntax hierarchy, no linear order, no phonological content

LINEARIZATION-1 = **Immobilization**

Morphosyntax hierarchy, adjacency, no phonological content

LINEARIZATION-2 = Vocabulary Insertion

Morphophonology no hierarchy, directed graph, phonological content

LINEARIZATION-3 = Serialization

Phonology no hierarchy, linear order, phonological string

Arguably, the *morphophonological module* and the *phonological module* are Markovian in nature, since there is no hierarchy. Between *morphosyntax* and *morphophonology* there must exist a dimensional flattening (in the terms of Krivochen, 2012b) algorithm, which transforms a hierarchical structure into a flat structure, without imposing extra structure. A phrase structure approach to vocabulary insertion and linearization, even though possible, is undesirable if a simpler solution is available. That is, in words of Lasnik & Uriagereka (2012), the "*inadequacy of powerful solutions to simple structuring*". Grammars which are high in the Chomsky Hierarchy are sometimes too complex for simple, Markovian structures; and the theory frequently falls in a diametrically opposite mistake as that pointed out in Chomsky (1957)¹³: Σ , F grammars (where Σ is a set of initial strings and F a set of post-style instruction formulae for rewriting) alone are inadequate for discontinuous dependencies, as in (27) (from Chomsky, 1957: 22):

1

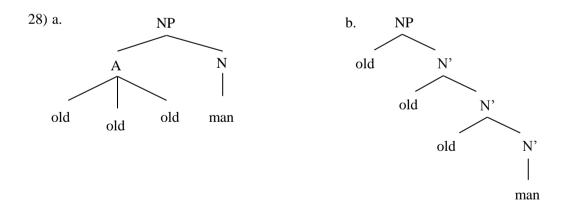
¹³ This is an essential point: the Hierarchy should probably be revisited, if the interpretation of "higher levels presupposing lower ones", since, should that be true, there would be no "additional structure" problem like that pointed out above. The mere idea of a mixed mind, looking for the simplest formalization for each particular type of cases, seems to call for interrelated study of the different formal languages, but by no means establishing an implicational hierarchy. A valid analogy, to the best of our knowledge, would be that of Euclidean, Hyperbolic, and Elliptical geometries. If we have a triangle whose inner angles sum 180 degrees, we will probably use Euclidean trigonometry to make calculations, not non-euclidean trigonometry: not because this makes calculations impossible (we well know it does not) or because there is a hierarchy of geometries, but because it is the simplest option for the problem in hand. Against this point of view, see Gallego (2007), who basically repeats Chomsky's case.

27) a. If S_1 , then S_2

b. Either S₃ or S₄

c. The man who said that S₅ is arriving today

The problem is described in terms of the "recursion-iteration" opposition in Chomsky's work. However, since "recursion" is an undefined term even today (see for example the Everett-Pesetsky debate about Piraha, mainly due to the lack of agreement on a criterion to determine the presence of recursion and the use, as synonyms, of 'recursion', 'embedding' and related terms in the critics), let us try to phrase the problem in less problematic terms. We agree with Chomsky in that there are great portions of human languages that cannot be appropriately described by means of finite state grammars, as those exemplified for English in (27). However, it would be a mistake to think that phrase structure grammars, either incorporating a transformational algorithm or not (e.g., HPSG, LFG, CG) can account for all constructions in all human languages. The reason, we argue (somehow following the line of reasoning of Lasnik & Uriagereka, 2012) is that there are naturally Markovian objects in natural languages which resist phrase structure description. We saw in previous sections that the Chomsky Hierarchy was sometimes too weak to account for (say) quantum phenomena: now, we add that it is sometimes too powerful insofar as natural languages are classified as phrase structure grammars plus a transformational component, with the computational and formal requirements this implies. Going back to our example (22), there are several ways in which one could represent the structure involved, we will just compare two:



If sisterhood imposes relations of scope (as c-command definitions lead us to assume, either in representational –Reinhart, 1976- or derivational –Epstein et. al. 1998- versions), then (28 b) is imposing too rich a structure for what is really a flat relation between elements, without any of them having scope over the others. A strict phrase structure model (e.g., Chomsky & Miller, 1963) is thus inadequate, we have to go one step below the Chomsky Hierarchy. Notice, incidentally, that (28 a) could be generated with a Σ , F grammar, where Σ = A and F = terminal strings (lexical items) but only

allowing F to be infinite (since there can be infinite instances of "old"), which is a trivial generative procedure, apart from computationally and biologically implausible. Formally, it would tell us nothing (as a non-trivial procedure must be restrictive enough to determine conditions of well-formation, in a Standard-Theory-like grammar), and empirically, it would generate too much. A Markovian representation, then, is not only a desirable scenario, but, as far as we can see, the only plausible one.

As regards mathematical modeling, it is to be noticed that a step-by-step derivational engine (be it Markovian or not) is modeled using difference equations, which allow us to calculate the state of the system at T_X as a function of the preceding terms T_{X-n} , T_{X-y} . The Fibonacci sequence dynamics that Uriagereka (1998, 2012) finds in clause structure, for instance, is an example of these kind of equations. For any term F of the sequence,

29)
$$F_n = F_{n-1} + F_{n-2}$$

If Fib is to be generated via an L-grammar of the kind Σ , F, however, it is not clear whether a difference equation could help in giving us the generative procedure used to get to a certain derivational point. This is particularly visible in the development of Phrase Structure Rules of the kind discussed in Chomsky (1957): unless we know that $S \to NP$, VP; given VP it is impossible to know how the system got there. Bottom-up models, on the other hand, could make better use of difference equations in developing generative algorithms which build the tree "from the bottom", independently of how many terms are involved in a *concatenation* relation.

Provided the notion of *frustration* we have introduced before actually applies to mental systems, as we believe, there would be an interesting tension arising here: the consideration of step-by-step derivational mechanisms within the mind seems to call for a difference equation modeling, but global tendencies, arising in complex systems with continuous time (that is, not chunked as we have done before) seems to call for a differential equation modeling. Consider a symbolic object derived via, say, *monotonic concatenation*. The step-by-step bottom-up derivation could be modeled using difference equations, but the overall pattern is that of a self-similar fractal: any syntactic object or arbitrary complexity, can be subordinated to another or establish with another a paratactic relation giving origin to a new object containing two complex units. Thus, if, according to Madrid (2011: 67), "a continuous dynamic system is chaotic if and only if there is a Poincaré section in which a discrete chaotic system can be defined" [our translation], it is highly likely the global tendency in linguistic (narrowing our scope down) computations is differential, whereas the on-line, local dynamics obey difference equations. The issue is very interesting and potentially revealing, and is the center of our current investigation.

We have briefly reviewed instances of Markovian objects within language, both in phonology and in the so-called "narrow syntax". Their presence was predicted by our model, if the mind actually displays geometrical frustrations in different sub-systems. This means that, just as a pure connectionist or purely modularist model do not accurately describe high-level and low-level processing (Carreiras, 1997: Chapter 4), providing arguments for mixed models which include connectionist networks for non-symbolic structures (being focused on interactive, multi-layered neural networks) and modular architectures for generative, uni-directional processes (as in Fodor's 1983 model, whose unidirectionality is shared by orthodox Chomskyan syntax by virtue of its syntacticocentrism, as Culicover & Jackendoff point out); the development of a mixed model, including different kinds of structures (Markovian, linear, chaotic, and quantum) seems to be a plausible road to take.

4. Conclusion

In this paper we have argued in favor of the existence of quantum processes in the mind, exemplifying with (but by no means limiting ourselves to) natural language. In our argumentation it became obvious that trying to subsume all computational processes in the mind to a single model (Markovian, phrase-structural, transformational, quantum) results in failure due, and that, just as it happens with neural networks, a mixed approach, distributing phenomena between different layers in the Chomsky Hierarchy, is at the same time more powerful and simpler. What is more, we have seen that the Chomsky Hierarchy (if it is to persist) is to be enriched with non-linear grammars, including chaotic and quantum phenomena. As Stapp (2009) puts it, quantum mechanics allow us to bridge the gap between "mind and matter" without the need to resort to stipulations in either side. We are well aware that there have been recent attempts to unify computational processes, manipulation of symbolic representations (for example, the "Turing program for linguistic theory" advocated for by Watumull, 2012; as well as the 'Flat Structure' proposal by Culicover & Jackendoff, 2005), but we doubt they can accommodate all the phenomena we have briefly presented and discussed here. If anything, the present work is a plea for mixed approaches and multidisciplinary interaction, focusing on language but without forgetting it is an integral part of the natural world and should not be studied in substantive isolation.

5. Bibliography

Ackerman, F., G. Stump & G. Webelhuth (2011) "Lexicalism, Periphrasis, and Implicative Morphology". In Borsley, R. & K. Börjars (Eds.) *Non-Transformational Syntax. Formal and Explicit Models of Grammar*. London, Blackwell. 325-358.

Aronoff, M. (1976) Word formation in generative grammar. Cambridge, MA: The MIT Press.

Baddeley, A. (2003) "Working Memory and Language: An Overview". *Journal of Communication Disorders* 36 (2003) 189–208

Binder, P. (2008) "Frustration in Complexity". Science 320. 322-323.

Boeckx, C. (2010) Defeating Lexicocentrism. Ms. ICREA/UAB. lingBuzz/001130

Borer, H. (2005) In Name Only: Structuring Sense vol I. Oxford: Oxford University Press.

(2009) Roots and categories. Talk presented at the 19th Colloquium on Generative Grammar, University of the Basque Country, April, 1-3 2009.

Carreiras, M. (1997) Descubriendo y procesando el lenguaje. Madrid: Trotta.

Chomsky, N. (1957) Syntactic Structures. The Hague, Mouton.

(1970) "Remarks on nominalization". In: Jacobs, Roderick & Peter Rosenbaum (eds.) *Readings in English transformational grammar.* Waltham, MA: Ginn. 184-221.

(1995) The Minimalist Program Cambridge, MA. MIT press.

(1998) Minimalist Inquiries. The Framework. *MIT Occasional Papers in Linguistics* 15.

(1999) Derivation by Phase. MIT Occasional Papers in Linguistics 18.

(2005a) On Phases. Ms. MIT.

(2005b) The Biolinguistic Perspective after 50 years. Sito Web dell'Accademia della Crusca - Aprile 2005. In *Quaderni del Dipartimento di Lingüística* (14), Firenze.

Collins, C. (2002) "Eliminating Labels". In *Derivation and Explanation in the Minimalist Program* (eds S. D. Epstein and T. D. Seely), Blackwell Publishing, 42-64.

Culicover, P. & R. Jackendoff (2005) Simpler Syntax. Oxford: OUP.

De Belder, M. (2011) *Roots and Affixes: Eliminating Lexical Categories from Syntax*. PhD Dissertation, Utrecht University.

De Belder, M. & J. van Craenenbroeck (2011) "How to merge a root". Ms., HUBrussel & Utrecht University.

D' Espósito, M. (2007) "From Cognitive to Neural Models of Working Memory". *Phil. Trans. R. Soc.* B 29 May 2007 vol. 362 no. 1481. 761-772

De Lancey, S. (2001) *Lectures on Functional Syntax*. Ms. University of Oregon. Available at: http://www.uoregon.edu/~delancey/sb/functional_syntax.doc

Dennet, D. (1991) Consciousness explained. Boston: Little, Brown and Company.

Einstein, A.; B. Podolsky & N. Rosen (1935). "Can Quantum-Mechanical Description of Physical Reality be Considered Complete?". *Physical Review* **47** (10): 777–780

Epstein, S., E. Groat, R. Kawashima & H. Kitahara (1998) *A Derivational Approach to Syntactic Relations*. Oxford University Press: Oxford.

Epstein, S. & T. D. Seely eds. (2002) Derivation and Explanation in the Minimalist Program. Oxford: Blackwell.

Fábregas, A. (2005) La definición de la categoría gramatical en una morfología orientada sintácticamente. PhD Dissertation, UAM.

Freeman W.J. (2000) *Neurodynamics*. An Exploration of Mesoscopic Brain Dynamics. London UK: Springer-Verlag.

Freeman, W. & G. Vitiello (2005) "Nonlinear brain dynamics as macroscopic manifestation of underlying many-body field dynamics". Ms. <u>arXiv:q-bio/0511037v1</u>

Fodor, J. (1975) The Language of Thought, Harvard University Press.

(1983) The Modularity of Mind: An Essay on Faculty Psychology. Cambridge, Mass.: MIT Press

Gallego, A. (2007) "Connectivity in Markovian Dependencies". *Proceedings of ConSOLE XIV*. 73-98.

Green, G. (2011) Elementary Principles of HPSG. In Borsley, R. & K. Börjars (Eds.) *Non-Transformational Syntax. Formal and Explicit Models of Grammar*. London, Blackwell. 9-53.

Greene, B. (1999) The Elegant Universe. New York, W. W. Norton.

Hale, K. & S. J. Keyser (1993) On argument structure and the lexical expression of Syntactic Relations. In *The view from Building 20: Essays in honor of Sylvain Bromberger*, ed. by Kenneth Hale and Samuel Jay Keyser, MIT Press.

(1997) The Basic Elements of Argument Structure. Ms. MIT. (adapted as Chapter 1 of Hale & Keyser, 2002)

(2002) Prolegomenon to a Theory of Argument Structure. Cambridge, Mass.: MIT Press.

Halle, M. & A. Marantz (1993) "Distributed Morphology and the pieces of Inflection". In: Hale, Kenneth & Samuel Jay Keyser (eds.) *The view from building 20*. Cambridge: MIT Press. 111-176.

Hauser, M.D., N. Chomsky & W.T. Fitch (2002) "The Faculty of Language: What Is It, Who Has It, and How It Evolve?" *Science* 298, (5598): 1569–79.

Idsardi, W. & E. Raimy (in press) "Three types of linearization and the temporal aspects of speech". In T. Biberauer and Ian Roberts (Editors) *Principles of linearization*. Berlin: Mouton de Gruyter.

Jackendoff, R. (2002) Foundations of Language. Oxford: OUP.

Kayne, R. (1994) The Antisymmetry of Syntax. Cambridge, Mass.: MIT.

Kitahara, H. (1997) *Elementary Operations and Optimal Derivations*. Cambridge Mass.: MIT Press Lasnik, H. (1999) *Minimalist Analysis*. Oxford: Blackwell.

(2011) "What Kind of Computing Device is the Human Language Faculty?". In Di Sciullo, A-M. & C. Boeckx (Eds.) *The Biolinguistic Enterprise*. Oxford: OUP. 354-65.

Lasnik, H. & J. Uriagereka (2012) "Structure". In R. Kempson, T. Fernando, and N. Asher (eds.) *Handbook of Philosophy of Science Volume 14: Philosophy of Linguistics*. Elsevier. 33-61.

Litt, A., C. Eliasmith, F. Kroon, S. Weinstein & P. Thagard (2006) "Is the Brain a Quantum Computer?" *Cognitive Science* XX (2006) 1–11.

Madrid, C. (2011) La mariposa y el tornado: teoría del Caos y cambio climático. Madrid, RBA.

Marantz, A. (1997) "No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon". In: Dimitriadis, Alexis et al. (eds.) *Proceedings of the 21st Annual Penn Linguistics Colloquium: Penn Working Papers in Linguistics 4.2*, 201-225.

Moro, A. (2000) Dynamic Antisymmetry. Cambridge, Mass.: MIT.

Penrose, R. (1997) "Physics and the mind". In M. Longair (Ed.), *The large, the small and the human mind*. Cambridge: Cambridge University Press. 93–143.

Pesetsky, D. & Torrego, E. (2007) "The syntax of valuation and the interpretability of features". In *Phrasal and Clausal Architecture. Syntactic Derivation and Interpretation*, S. Karimi et. al. (eds.). Amsterdam: John Benjamins. 262-294.

Pylyshyn, Z. W. (2007) *Things and places: How the mind connects with the perceptual world* (2004 Jean Nicod Lectures). Cambridge, MA: MIT Press.

Reinhart, T. (1976) The Syntactic Domain of Anaphora. PhD Dissertation, MIT.

Schrödinger, E. (1944) What is Life? Cambridge, Mass.: CUP.

Sola, J. (1996) "Morphology and word order in Germanic languages". In Minimal Ideas: Syntactic

Studies in the Minimalist Framework, W. Abraham et al.(eds.). Amsterdam: John

Benjamins. 217-251.

Stapp, H. (2009) Mind, Matter and Quantum Mechanics. Springer.

Talmy, L. (2000) Toward a cognitive semantics. Cambridge, Mass.: MIT.

(2007) "The relation of Grammar to Cognition". In Evans, V., B. Bergen & J. Zinken (eds.) *The Cognitive Linguistics Reader*. London: Equinox. 481-544.

Uriagereka, J. (1998) Rhyme and Reason. MIT Press.

(1999) Multiple Spell-Out. In N. Hornstein & S. Epstein (eds.), *Working Minimalism*, Cambdridge (Mass.), MIT Press, 251-282.

(2002) Multiple Spell-Out. In Uriagereka, ed. *Derivations: Exploring the Dynamics of Syntax*. London, Routledge.

(2005) A Markovian Syntax for Adjuncts. Ms. UMD.

(2008) Syntactic Anchors: On Semantic Restructuring. Cambridge: CUP.

(2011) A Sketch of the Grammar in Non-Classical Conditions. Ms. UMD.

(2012) Spell-Out and the Minimalist Program. Oxford: OUP.

Vergnaud, J-R. (1977) Letter to Noam Chomsky & Howard Lasnik re. their manuscript "Filters and Control". Ms. http://ling.auf.net/lingbuzz/000461

Vitiello, G. (2001) My Double Unveiled. Amsterdam, John Benjamins.

Watumull, J. (2012) A Turing Program for Linguistic Theory. In *Biolinguistics*, 6.2. 222-245.

Williams, E. & A. M. Di Sciullo (1987) On the definition of word. Cambridge, MA: The MIT Press.

Wurmbrand, S. (2013) "The Merge Condition: A syntactic approach to selection". To appear in *Minimalism and Beyond: Radicalizing the interfaces*, ed. by P. Kosta, L. Schürcks, S. Franks, and T. Radeva-Bork. Amsterdam: John Benjamins