A Frustrated Mind

Diego Gabriel Krivochen

University of Reading, CINN

diegokrivochen@hotmail.com

Abstract:

In this work we develop the thesis that the concept of dynamical frustration (Binder, 2008), involving the tension between conflicting tendencies in a physical system, is essential when analyzing the cognitive-computational bases of language. We will argue that there are two kinds of frustrations interplaying, in principle as independent theses: an architectural global frustration relating nonlinear, phrase-structural, and finite-state processes in general computations; and a local derivational frustration relating semantic and phonological cycles within language processing. After developing and analyzing each frustration, with particular focus on human language, we will briefly explore the consequences these frustrations have on the "design" of mental faculties, and the impact our theory has on the Minimalist notion of perfection in language design.

Keywords: Dynamical Frustration; interactive computation; Chomsky Hierarchy; Language

1. *Introduction: what is a "dynamical frustration"?*

The introduction of the concept of *frustration* needs a prior analysis of the concept of *complex systems*. A complex system, to summarize Boccara (2002), is a set of elements and relations, such that the following properties arise:

- *Emergence*
- *Self-Organization*

Let us analyze each briefly. The concept of *emergence* has been around since Aristotle, and it has been used in life sciences (e.g., Corning, 2002) as well as economics (e.g., Goldstein, 1999) and physics. These uses have a common core: the behavior of the system cannot be predicted as a function of its interacting elements. In other words, complex patterns arise out of a multiplicity of simple interactions (see Hofstadter, 1979 for examples and discussion centered in formal systems, music, and the visual capacity). Consider Boccara's characterization of complex systems as having the following characteristics (2002: 3):

- They consist of a large number of interacting *agents*
- They exhibit *emergence*
- The emergent behavior does not arise from the existence of an external *controller*

Notice that the use of the word "agent" does not presuppose volition: Boccara himself analyses the stock market as a self-organizing complex system without taking the intentions of the stock buyers-sellers into account.

In natural language, we have a certain number –potentially unlimited since coinage processes are productive- of elements to manipulate (lexical items), the characteristics and behavior of the collective structure does not directly follows from the individual characteristics of the interacting elements (that is to say, the *meaning* of a syntactic structure is not only in the elements, but also in the way they organize, something that cannot be predicted from the array of elements alone), what is called *emergence*;[1] and finally, this emergent behavior is not the result of the existence of a *central processor*, as would be the case in traditional Fodorian modularism (Fodor, 1983). Crucially, emergence in language is a property that can be hypothesized to follow from its very architecture as a cognitive system: natural human language can be formally defined as the interaction (or, set-theoretically speaking, the *intersection*) between a generative algorithm and two interfaces, without it being possible that the separate characteristics of any of them determines the behavior of the system as a whole. Conceptual structures, for instance (resulting from the interface between generation and semantics) have different properties from language: they are not restricted by the need to externalize structure (phonological externalization being essentially finite-state, see Uriagereka, 2012; Idsardi & Raimy, in press), and therefore different kinds of dependencies between conceptual elements can be established from those possible within human language. In one way or another, then, language can (and, we claim here, should) be seen as a complex system.

Boccara relates *emergence* in complex systems to the absence of a central controller (something which would not arise in a Fodorian modularist model, but would be expected in a Massive Modular mind: see Carruthers, 2006, particularly Chapter 1). In this work, we will work with the concept of *strong emergence*, which implies that a property cannot be irreduced to the individual components of the system. In other words, *strong emergence* arises as a consequence of the way elements interact in what we will argue is essentially a syntactic structure, not of the nature of the elements themselves. This is essential for our proposal in this paper, since the theory of how elements interact in linguistic derivations is precisely the theory of *syntax*, narrowing the scope of "syntax" to language in the present work exclusively for expository purposes. At this respect, it is useful to compare our perspective on what 'syntax' is with that of 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, in contrast, and in general within our theory, "syntax" is used in a sense broader than that of Culicover and Jackendoff, for two main reasons: to begin with, there is no compelling evidence that the "syntactic mechanisms" taken alone (without considering the elements involved, just the combinatory algorithm, which is in our proposal 'blind' to the characteristics of the elements it manipulates, thus being able to generate complexity all through the mind) vary from one system to another, except that the units affect the algorithm, in case such process actually happens (see Krivochen & Mathiasen, 2012 for discussion of music and mathematics, as well as numerals within language); and also, an adequately wide formalization of syntactic mechanisms can reveal deep facts

[1] Interestingly, the concept of *emergence* can also be applied to *locality principles*: "*Emergent properties* are large-scale effects of a system of locally interacting agents that are often surprising and hard to predict" Boccara (2002: 23). If language displays chaotic characteristics, it is to be expected that it also displays locality effects, since locality is, apparently, a condition for complex systems.

about the structure of more than a single system (including, for example, the structure of protein and DNA folding, see e.g. Searls, 2002).

The concept of self-organization is also independent of composing units. This means that, while the system as a whole may display a certain behavior (e.g., asymptotical behavior towards a point or an attractor), no individual in the system has consciousness of this global characteristic, which means that self-organization is not driven by a "leader" organism in biological systems or by a certain element present in a derivation in mathematical or linguistic structures. This means that we can only appreciate this tendency towards order if we, as observers, are outside the system (even if only theoretically, as in the case of social or population dynamics). Moreover, the concept of self-organization is highly dependent of that of *time*: a set of elements initially distributed randomly adopts a structured configuration in a process that, for expository reasons, we will consider discrete (i.e., varying the T dimension step-by-step, T_n, T_{n+1}, and so on): since we will work with a generative operation applying successively (monotonically) to *n* objects, a continuous approach to the derivational diachrony would make the exposition confusing. To pursue the characterization of self-organization, it is important to mention the proposal by Bak & Paczuski (1995), who summarize much previous discussion (e.g., Gould, 1989), and introduce the concept of critical point -of major importance in chaos theory and nonlinear systems- in self-organization. Given a difference equation (a specific type of recurrence relation, recursively defining a sequence in which each term is a function of the previous one, quite in the line of the step-by-step derivational procedure we advocate), there is a critic value for a variable, which determines the arise of chaotic behavior in a system. For instance, in Feigenbaum's (1975) system, the binary-branching graph in (1), where *r* is the critical point, behaves chaotically after $r \approx 3.6$

1) $f(x) = rx(1-x)$

The graph is shown in Figure I:

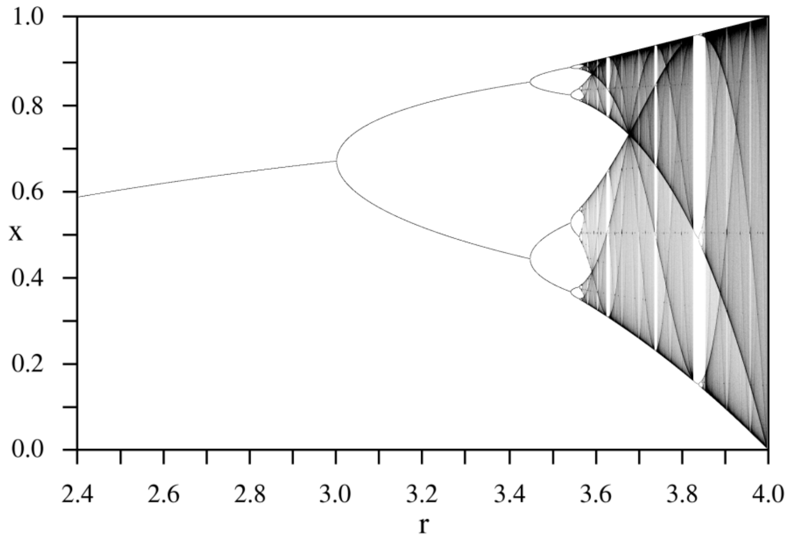


Figure I: Feigenbaum's logistic map

Notice that there are zones of binarity after chaos, which we will show represent derivational *cycles* (also commonly referred to as *domains*). Along the lines of Uriagereka (2011)—although there are considerable differences between his approach and ours— we have characterized a linguistic derivation as a finite set of applications of a structure-building algorithm, in which entropy values fluctuate after each application (Krivochen, 2013), tending towards minimization of entropy within each derivational *cycle* in a semantically-driven syntax. We will return to the concept of *cycle* below, when dealing with the problem of structure externalization.[2] These cycles of alternating entropy, where structure building cumulatively reduces entropy both locally and globally (for a cycle and for a whole derivation) as we derive a well-formed formula in a formal system (in language, a sentence in a natural language L, abstracting performance factors solely because of methodology), we will argue, are closely connected to considerations of locality over dependencies. These locality considerations rule, for instance, the interface-optimal derivational distance between an anaphoric element and its antecedent (as in 2), superiority effects in English Wh-interrogatives (in 3), and variations in quantifier scope depending on the closest argument at the level of Logical Form (in 4, following the theory of Hornstein, 1995), to mention but a few:

2) a. Mary_i likes a picture of herself_i
   b. Mary_i said [that Anna_j likes a picture of $\text{herself}_{*i/j}$]
3) a. Who wants what?
   b. *What who wants?
4) a. Every man loves a woman
   Logical Form: $\forall(x)\ \exists(y)$ (x is a man & y is a woman & x loves y)
   $\exists(x)\ \forall(y)$ (x is a man & y is a woman & x loves y)

As argued by transformationalist generativists and non-transformational generativists alike (see Patel-Grosz, 2012; Müller, 2011; Boeckx, 2008 on the one hand, and Stroik, 2009; Sag, 2007; Green, 2011 on the other, for recent examples of both positions), *locality* on the establishment of dependencies between constituents is a pervasive property of syntactic structures. In brief, the orthodox position has, in turn, two aspects: the *locality-as-impenetrability* concept of Chomsky (1986, 1999, 2008), according to which there are certain syntactic objects, defined beforehand, which are impenetrable to external operations (so-called "barriers" back in the '80s, now dubbed "phases"[3]); and the *locality-as-non-interveniency* proposal made by Rizzi (1990, 2009), which is best expressed in the form of the Relativized Minimality Principle, formulated as follows (Rizzi, 2009: 162):

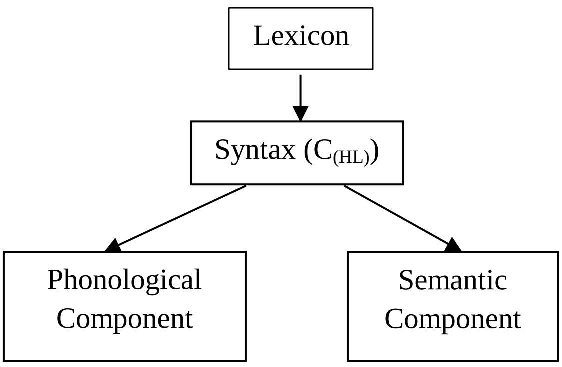
> …*in a configuration like the following:*
> . . . *X*. . . *Z* . . . *Y*. . .
> *no local relation can hold between X and Y across an intervening element Z, if Z is of the same structural type, given the appropriate typology of elements and positions, as X.*

[2] David Medeiros (p.c.) has commented in this respect that "My uninformed intuition strongly suggests that, *qua* natural system, it should rather maximize entropy. I had read something recently about extending this to the idea that natural systems tend to maximize *future* entropy, (which turns out to make different and better predictions in some domains)." We think this objection is too important to be left aside (although we lack access to the predictions Medeiros mentions). A summary of the extensive case we have made in Krivochen (2013) would be as follows: Assume a type-array {a, b, c, d}. The first derivational step is the most entropic, since any option is equally likely to enter the derivational workspace. Suppose {a} is selected. The derivational options are reduced now, since not all combinations are readable by the semantic interface, which I maintain drives the derivation. In this sense, if Merge is semantically driven, it minimizes entropy in a local level, until we reach a derivational situation in which, again, all options are equally possible. There, we have maximum entropy, again, and incidentally, we can define phases *qua* periodic measurements of entropy (the interpretative interfaces accessing the syntactic workspace), as "cycles" with entropic peaks and sines. Our system, then, is not uniformly but *locally* "entropy-reducing". The notion of *local cycle* is of central importance in our model.

[3] See Boeckx & Grohmann (2004) on this issue, as the similarities between *barriers* and *phases* are too obvious to be ignored.

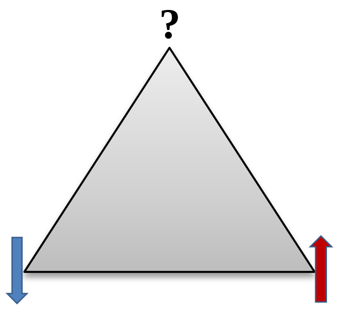
Needless to say, there are complications with both approaches. Chomsky's theory requires stipulative definitions of what constitutes an impenetrable object and what does not, which has been already challenged (e.g., by Uriagereka, 2002a, who proposes to define local domains according to phonological / linearization requirements, segmenting complex representations of the kind {{a, b}, {c, d}} in sub-chunks of finite-state compatible objects, {a, b} and {c, d}). Rizzi's theory requires a rich typology of elements and positions (something the author himself recognizes), including heads, non-heads, A(rgumental), and A' (non-argumental) positions; which makes the theory inelegant and biologically implausible (no model so far has even attempted to find neurocognitive correlates for those required notions, not to mention additional complications if feature-valuation/checking considerations enter the game, which are quite central in the current Minimalist system, see, e.g., Epstein et. al., 2010; Den Dikken, 2014; Putnam & Fábregas, 2014). Both share the characteristic of being syntacticocentric (in terms of Culicover & Jackendoff, 2005), which amounts to saying that the focus is set on the syntactic component, isolated substantively *and* methodologically from semantics and phonology, and constraints on well-formed formulae are exclusively syntactic in nature. The syntactic component (referred to as the Computational System for Human Language $C_{(HL)}$) then handles (or, technically, "transfers") chunks of information (proper subparts of the overall derivation; subtrees, in a more visual way) to the external systems in charge of semantics and more generally, conceptual interpretation (the so-called Conceptual-Intentional system C-I) and morpho-phonology (the Sensory-Motor system S-M), the directionality of this process being represented by arrows instead of simple lines in (5). The orthodox generative architecture for the language faculty can be schematized as follows (the so-called Y-model), where arrows indicate the unidirectional flow of information:

5)



What is the role of dynamical frustrations in this discussion, focusing on language? So far we have been discussing some properties of dynamical systems, analyzed through difference equations. Binder (2008) claims that the main characteristic of dynamical systems (which sometimes appear to be very different in a surface level, take heartbeat rate and climatic conditions as examples) is that *they display a fundamental tension between global and local tendencies*, which is called a *scale frustration*, a specific instantiation of a more general phenomenon known as a *dynamical frustration*. These frustrations arise, for example, in a triangle lattice under the condition that all three be antialigned (that is, aligned in an opposite direction) with respect to one another: three neighboring spins cannot be pairwise antialigned, and thus a frustration arises (Moessner & Ramirez, 2006: 25-26; Binder, 2008: 322). Let us graph the situation:

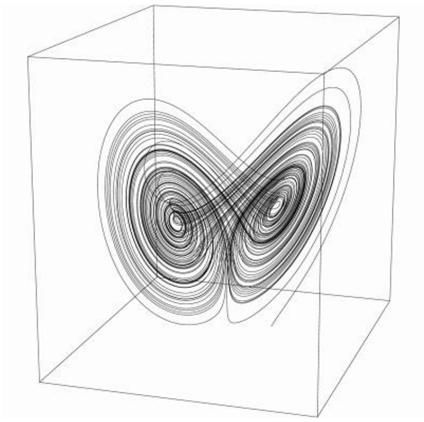
6)



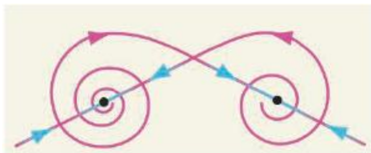
Binder goes even further, positing that a complex system where no frustration arises will "*either settle to equilibrium or grow without bounds*" (2008: 322). Equilibrium, Binder's sense, is not *balance*, which could be, in principle, desirable: we are not referring to an optimal organization of matter or distribution of forces, but to complete lack of activity (for instance, a pendulum, which, after a while, stops moving altogether).

Crucially, a system can display opposing tendencies at different scales: local and global tendencies may result in a "lattice triangle situation" like the one graphed in (6). A relevant example is the well-known Lorenz attractor (7a, b) (which describes the behavior of a chaotic system), where there are opposing clockwise and counter-clockwise tendencies in the 3-D phase space (7 b is taken from Binder, 2008: 322):

7) a)



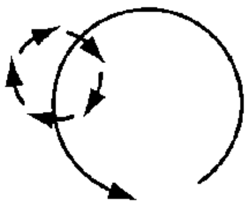
b)



In the 2-D figure (7 b) we can see opposing tendencies between "*stretching out and folding*", in a resemblance of centripetal and centrifugal forces pulling the attractor in opposite directions. As the attractor "grows" in each dimension, centrifugal forces can be informally said to prevail. This kind of frustration, exemplified by chaotic systems displaying dynamics like those graphed in (7), is called a *geometrical frustration*.

There is yet another kind of frustration, which arises when global and local tendencies are opposing in nature: so-called *scale frustration*. Binder's example is a counter-clockwise global tendency with local clockwise cycles, as in (8):

8)



As we have argued in Krivochen (2015a, b), this kind of frustration may in fact be essential when considering the cognitive processes that take place in both the production and interpretation of natural language. What is more, we have argued in those work, and will argue here (further following Uriagereka, 2012: Chapter 6; 2014), that the concept of frustration is also of key importance when trying to situate language in relation to other cognitive capacities. If language is indeed a locally chaotic system,[4] with derivational cycles displaying locally decreasing entropy as more structure is built and thus there is more certainty for a parser with respect to where relevant interpretations are to be found in a semantic phase space, for example (as put forth in Krivochen, 2013; see also Uriagereka, 2011 for a different position, although stemming from the same assumptions), then the *geometrical frustration* illustrated in (6) is relevant, insofar as the mathematics associated to a Lorenz attractor, or Feigenbaum's equation could prove useful for modeling linguistic derivations (from conceptual structures to externalized sound waves) and the cognitive capacities that licenses them. On a different scale, the interplay between different components in language (namely, syntax, semantics, and phonology), and also between language and other cognitive capacities (and the respective computational procedures underlying them), including the so-called 'number sense' (Dehaene, 1997), the music capacity, among others (see Hofstadter, 1979 for an overview), would require a model in line with the concept of *scale frustration*, to account for global and local effects in different components and their interplay (using Minimalist terms, their *interfaces*).

The notion of "scale" in frustration is essential for our argumentation. We will first work with two different kinds of frustrations, at different levels and each having local and global tendencies. Then we will attempt to unify them in a single kind of frustration.

2. A frustrated mind

We will argue, following the line of previous works, that there are two kinds of frustrations at two different scales to be taken into account in cognitive studies:

- *Architectural frustration*
- *Derivational frustration*

In the following subsections we will analyze each in some detail.

2.1 *Architectural frustration: is the mind computationally uniform?*

In Krivochen (2015a), somehow following Uriagereka (2012), we have put forth the hypothesis that there is a fundamental tension in the human mind: a global tendency towards interactive computation (Goldin & Wegner, 2005, 2007), going beyond the function-based computational claims related to the Church-Turing thesis CTT (see Watumull, 2012 for an attempt to apply CTT to natural language); and an opposing, local tendency towards periodic strictly finite-state cycles. What is more, the *locus* of finite-state processes can quite uniformly be found whenever symbolic structures are to be provided phonological matrices in order to be externalized (i.e., when the Sensory-Motor systems come into play). The process of linearization of structure and their externalization is computationally closed, linear, and representationally local (meaning, it targets minimal structural domains, as soon as a representation is a suitable input for the linearization algorithm). In contrast, there are global symbolic structures whose interpretation requires information retrieval from different sources *during* the

[4] By "locally chaotic" we mean not only that the machinery is essentially derivational, but that the derivation, as it proceeds step-by-step, displays entropic behavior, in which entropy varies depending on which element enters the derivational space, and the kind of syntactic object thereby generated.

computation: we argue that an interactive view of computation is preferable in those cases, rather than trying to adapt or expand the limits of function-based computation. We will call this tension, being pervasive to more than one mental capacity, an *architectural frustration*, as it underlies the interplay between faculties. As such, it is wider in scope than the *derivational frustration*, which belongs to a lower or narrower scale.

To put our inquiry in context, the Cognitive Revolution of the ‘50s, which had a great impact on theoretical linguistics, brought along 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-computable models of the mind, and even a “Turing Program for Linguistic Theory” (Watumull, 2012), based on the claim that cognition can be modeled *uniformly* within an unlimited-memory, recursively enumerable grammar (see Uriagereka, 2012, Chapter 6 for general discussion; Krivochen, forthcoming b for discussion of the empirical limitations of a uniform cognitive template for phrase structure). The belief that linguistic computation is function-based and thus falls within Turing limits is already a commonplace in (bio)linguistic theorizing. From the arguments against finite-state models in Chomsky (1957, 1959) to the latest developments in Watumull et. al. (2014), encompassing Joshi (1985), Galley & McKeown (2006), Yli-Jyra, Kornai & Sakarovitch (2011), Uriagereka (2008, 2012) -to name but a few-; linguistic computation has been hypothesized to be exhaustively described by means of a Turing machine. While not all aforementioned authors agree with respect to *which* computational procedure is the most accurate (for example, Joshi’s 1985 Tree Adjoining Grammar can be implemented by means of a Push Down Automaton+, whose memory is not strictly *last-in-the-stack* limited), they are all ultimately expressible in terms of a deterministic Turing Machine (DTM), given Chomsky’s theorem (1959: 143):

9) *Theorem 1. For both grammars and languages, type 0 $\supseteq$ type 1 $\supseteq$ type 2 $\supseteq$ type 3.*

Provided that

10) Type 3: Finite state models

    Type 2: Context free grammars

    Type 1: Context sensitive grammars

    Type 0: Turing machines

In a word, as the standard CTT assumes,

11) *TMs can compute any effective (partially recursive) function over naturals* [natural numbers] *(strings)* (Goldin & Wegner, 2007: 16).

However, it is not clear that this kind of computational device is the most appropriate for all linguistic computation, as it is based on the concept of *function*. For starters, we must make explicit what we include within the term ‘function’. The definition we adopt is the following (based on Falcade et. al., 2004; Youschkevitch, 1976/1977: 39; May, 1962, among others)

> *A function is a relation between a set of inputs and a set of permissible outputs with the property that each input is related to exactly one output.*

That is, a function establishes a univocal relation between input and output (or sets thereof). A function-based machine provided with a suitable (possibly finite) input and a finite set of rules for its manipulation, can generate a (possibly finite) output (in polynomial time, if both sets are finite). Consider, for example:

12) $f(x) = x^2$

This function relates each value of x to its square x^2 by means of a definite rule, 'multiply x by itself'. The alphabet is the set of integers $\mathbb{Z}$, and provided an infinite tape as in a TM, the function can go on and on, until halting is stipulated by a rule. Crucially, neither the alphabet nor the rules can be changed in the middle of the process, nor can information from other systems influence the process. In and by itself, there is nothing objectionable to this view of computation, if it were not for the fact that it is a common procedure to equate *computation* to *computation of functions*, as David (1958) eloquently puts it in his introduction to a classic textbook:

> *This book is an introduction to the theory of computability and non-computability, usually referred to as the theory of recursive functions.*

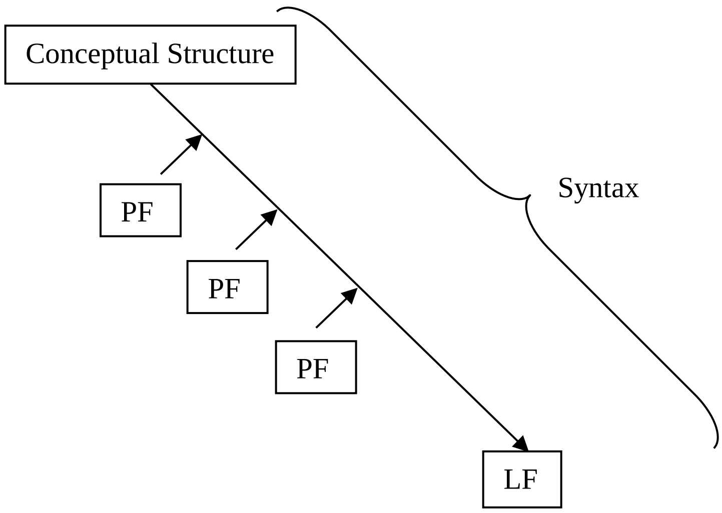
We are thus faced with *two* problems, not just one.

The CTT is a claim that is to be separated, as Goldin & Wegner (2007) point out, from the assumption that *computation* equals *computation of functions*. This has far-reaching consequences, for even those who have tried to increase the power of TMs in order to compute more complex problems (e.g., Deutsch, 1985, who holds a physical view of TMs and expands the concept to a Universal Quantum TM; Santos, 1969 and his analysis of *probabilistic* TMs; among many others) have come up with algorithmic extensions that suffer from one of two types of problem: either they are formally *equivalent* to a deterministic TM (such extensions employ procedures such as increasing the number of tapes, changing the alphabet, and so on), and have yielded little, at the cost of numerous extra assumptions; or they are quite simply *not TMs* in the sense of Turing (1936).

When dealing with language as a mental capacity, neurologically rooted, some theories adopt the CTT stance explicitly. Thus, for example, the Minimalist Program (Chomsky, 1995, 2000 et. seq.) is a function-based approach, in which the generative operation Merge is a function taking two arguments (Chomsky, 2013; and much related work), the output being determined linearly from the input. Earlier versions of the generative theory (Chomsky, 1957; 1965, for example) are also explicitly based upon the concept of *function*, and the Government and Binding (GB) model, which was developed during the 1980s, embraced a form of modularity which prevented bidirectional informational flow and interaction between different modules except at certain points, or levels of representation (D-Structure, S-Structure, Logical Form, Phonetic Form).

The assumptions about the complex, nonlinear character of language, and the existence of conflicting tensions at the root of cognitive capacities we have made here are related (even if in a non-necessary way) to a proposal about the architecture of the cognitive system underlying language production and comprehension, and the mathematics necessary to model it, which we will present here. The architecture we assume is the following, where cycles are locally PF (Phonetic Form)-driven, as in Uriagereka's 2002a Multiple Spell-Out (MSO) model, but the overall system is LF (Logical Form)-driven, primarily by principles of global *conservation* (Lasnik and Uriagereka, 2005: 53; Krivochen, 2011: 53); thus generating a tension (a discontinuity or *hiatus*, in Uriagereka's terms) between computational requirements:

13)



In the architecture in (13), unquestionably inspired by the MSO model, the directionality is even more obvious: there is a single arrow from CS to LF, the global path where conservation is the major force; several arrows *from* PF *to* the syntactic workspace indicate multiple accesses of the PF component to identify and 'co-opt' linearizable chunks of structure. Now, what exactly are we conserving? It depends on what we think the conceptual structure or 'conceptual soup' is, the information it contains, and the format in which that information is structured, including the non-trivial issue of its dimensionality and its topological characteristics. It is to be highlighted that, since we consider 'syntax' is pervasive in all cognitive processes, as essentially relational, we have it all the way through, not emerging from the frustration. This is the first difference from the 'traditional' CLASH model (Uriagereka, 2012; 2014) in which the concept of *dynamical frustration* is implemented: our conception of 'syntax' (and 'phrase structure', among other related notions) is not only wider, but also more dynamic (as we reject the imposition of *structural uniformity*, i.e., a common template for all phrase structure, along with Culicover and Jackendoff, 2005: 7, ff.). If, considering a formal system, an element in the alphabet has the potential to yield many outcomes, determined by syntactic context and interface reading, prior to its insertion in a derivational context via a simple operation of symbolic *concatenation* (Krivochen, 2015a) that blindly and freely puts things together (without any further specification like *labels*), then we eliminate all kinds of diacritics (including categorial, case, theta, φ- (person / number), Time, Aspect, and Modality specifications), which are hard to avoid in lexicalist theories (due to the Lexical Integrity Principle, which asserts the opacity of lexical representations for syntactic operations thus separating the domains of the Lexicon and the "narrow syntax", see Ackerman et. al., 2011: 326; and Carruthers, 2006: 5-7 for some discussion related to the concept of *inaccessibility*, which applies to the lexicalist Lexicon). Such diacritics are also hard to avoid in late-phonology-insertion theories like Distributed Morphology (which "distributes" lexical information in three components, or "Lists" A, B, and C; consisting of formal features, phonological matrices, and encyclopedic knowledge respectively), where specific elements dubbed *categorizers* are needed to specify the categorial status of a root, and thus its distribution and selectional properties (Halle & Marantz, 1993). The reason we reject these elements is that their nature is, at best, unclear: are they formal elements, introduced in a derivation to motivate transformations and then be eliminated (as in Chomsky's 1995 et seq. proposal)? Or are they semantic or phonological (substantive, more generally) in nature? How are their values specified in a lexicon? If we examine each one of those diacritics in detail, we can simplify the alphabet of our formal theory by eliminating some, and establishing the contextual mechanisms by means of which other arise in the course a derivation (but not before) following global conservation principles.

By abandoning a procrustean approach to structure building, the "computational system" which generates symbolic structures is also drastically simplified: on the one hand, the 2-D and binary-branching limitations for phrase structure X' theory imposes are eliminated, as the model is *semantically* oriented (as opposed to mainstream Minimalism's 'syntacticocentrism', see Culicover and Jackendoff, 2005: 17), and there are neither proofs that semantic structures are bound by phrase structural requirements, nor theoretical reasons (beyond excessive simplification) to assume a strong uniformity among different cognitive domains (Culicover and Jackendoff, 2005: 17, 20; Krivochen, 2015a, b for an analysis of structure building and structure mapping under these assumptions, respectively).

If the so-called interface systems can access the syntactic workspace dynamically and they are actually the ones that drive structure building and mapping processes, pre-defined domains are also eliminated (Cf. Chomsky's 2000 phases and, to a lesser extent, Grohmann's 2003 Prolific Domains, both of which delimit relevant derivational chunks for the application of operations): each structural chunk is optimal and *optimality* is, we argue, a relational notion related to the resolution of a tension, not an aprioristically condition imposed over structure or representations. The scale is, as we said, essential to this claim. While global semantic computations seem to display properties beyond strict α-machine-like (automatic machine, see Turing, 1936: 232) Turing-computability (including interaction with non-linguistic information sources: contextual, cultural…; as well as selective backtracking, and the possibility of tampering with a derivation if things do not seem to go right, along the lines of Townsend and Bever's, 2001 *Analysis by Synthesis*), local processes are often better expressed by computationally simpler models, sometimes, the only way to explain or model certain processes is to assume there is another kind of computational procedure applying, otherwise, we would be assigning a certain representation "more structure" than it really has (for example, while phrase structure grammars are adequate to capture the properties of adjectival modification and scope, they are too powerful to deal with simple adjective iteration, when there is no scope relation between the terms involved; see Krivochen, 2015a; also Lasnik, 2011; Uriagereka, 2008: Chapter 6, for different, yet related proposals about the Markovian nature of iteration), thus begging the question whether that really is 'the best theory' (taking the expression from Postal, 1974). The conflict between different requirements from different systems, as well as an overall will to 'keep it simple' (as simple *as possible*) arises at each derivational point where Transfer (PF) applies. Taking into account Idsardi & Raimy's (in press) decomposition of Spell-Out in three steps, *n*-dimensional phrase markers must undergo a dynamic process that makes them finite-state compatible, which is customarily referred to as 'Spell-Out'. These authors distinguish three "modules" of linearization, with different characteristics (Idsardi & Raimy, in press: 3):

14) ***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 computationally finite-state (displaying the Markov property) in nature, since there is no hierarchy in linearized strings and no memory is required. Between *morphosyntax* and *morphophonology* there must exist a dimensional flattening (in the terms of Krivochen, in preparation; section 3.1 below; see Fodor, 2013: 217 for relevant discussion) algorithm, which transforms a hierarchical structure into a flat structure without imposing *extra* structure but following a strict *conservation principle* (Krivochen, 2011; Lasnik, Uriagereka & Boeckx, 2005: 53 for discussion of Conservation Principles under orthodox Minimalist assumptions; also Martin and Uriagereka, 2014; Uriagereka, 2008: 14, ff., 99). A phrase structure approach to vocabulary insertion (assuming a separationist approach in which terminals receive phonological matrices at Spell-Out, and not before, see Halle & Marantz, 1993 for discussion and a "late-insertion" proposal) and linearization, even though possible, is undesirable if a simpler solution is available; moreover, it has to be empirically tested. Taking into consideration the architecture sketched in (13), this architectural frustration between global high-dimensional processes and local finite-state processes takes place at each point of contact between the CS-LF arrow and the PF-leading arrows. More generally, and even if there is no language involved, conceptual structures, with their underspecification and *n*-dimensional character, are better expressed by means of non-linear mathematical models (in Krivochen, in preparation, we have used ordinary and partial differential equations), whereas those might be too powerful for the strictly linear character of externalized strings. An explicit model of how we get from high-dimensional structures to linearized strings is still under development (Krivochen, in preparation; see section 3.1 for an outline of some technical aspects of the model).

Needless to say, the fact that the *rules* that generate (in the strong sense) a string or a structure are finite-state does not mean that the (weakly) generated object is actually finite-state. As Belk & Neeleman (2013) point out,

> (…) *since the early 1980s, there has been an attempt in syntactic theory to abandon rules that have a linear component in favour of rules that only refer to dominance and labeling. Part of the motivation for this trend is the pervasive necessity of reference to structure in the description of linguistic phenomena.* (2013: 1)

When we refer to the linear character of a representation, or the finite-state character of a derivation (or more generally, any object of inquiry), we are hypothesizing about *the object itself*, regardless of the character of the formal apparatus used to model the object. To give an example, an L-grammar can be used to generate non-linear sequences, if it is complemented by an appropriate theory of node multidominance (for linguistic details, see Citko, 2011), or, more generally, a specification about the Hamiltonian or non-Hamiltonian character of a particular graph. *It seems that uniform models are inherently insufficient, not because of limitations of the models themselves, but because of the mixed nature of the objects they model* (a theory we tested in Krivochen, 2015a with respect to structure building). This mixed nature gives rise to our *architectural frustration*.

2.2 *Derivational Frustration: Semantic or Phonological Preeminence?*

The architecture in (13) makes a strong statement with respect to the status of cognitive interfaces, which in turn might be interpreted as a prediction: semantics structures the linguistic computational system, the Faculty of Language in the Narrow sense in the sense of Hauser, Chomsky & Fitch (2002) (HCF hitherto). However, in opposition to HCF, we do not claim that *concatenation* is language specific: conceptual structures, which might or might not be linguistically instantiated, are structured symbolic representations, and therefore syntactic in our sense. In the vein of linguists like Jackendoff (1983, 2002, 2010)[5], Culicover & Jackendoff (2005), and following the way in which D-Structure is interpreted in Uriagereka (2008, 2012); the conceptual structure we propose is syntactically assembled, displaying growing complexity, recursion in its generative procedure, and other formal properties of syntactic structure. Moss et. al. (2007) and Taylor et. al. (2011) propose a (cognitive-neurological) model for conceptual structure, which is focused on the inner structure of "concepts" (what some linguists, particularly adopting a Distributed Morphology approach, would call *roots*). The conceptual structure model proposed by Taylor et. al. (2011) differs from ours insofar as they take the word (or, rather, the "concept") as already formed, whereas we consider that the distinctive features they require for object identification (what we would call "reference", from a semantic-linguistic point of view, quite in a Strawsonean sense, see Strawson, 1950) are *not* part of the root itself, but are actually provided by procedural material taking the grammatical form of functional categories (tense, causativity, location) that limits the intensional / extensional scope of the semantic substance, pre-linguistic and possibly universal, conveyed by the root alone.

We pursue a complementary model, focused on how those "concepts" are manipulated by syntactic means, to form structures that might, or might not, be expressed linguistically (but, for instance, through action). This implies that the elements manipulated to create conceptual structures (in the more complex sense we talk about) must be underspecified enough to be manipulated by more than one system. We agree with the claim in Caramazza & Mahon (2006) that conceptual knowledge must exist independently of the modality-specific input and output systems (e.g., perceptual and articulatory systems). The organization of both input and output systems is assumed to be determined innately, and enriched with evolutionarily relevant conceptual categories according to which to organize information, cultural and other developmental factors also entering the picture (but we will not deal with mind-external factors here). With a more narrow linguistic emphasis, Moss et. al. (2007) claim that

> (…) *conceptual representations form an interface between lexical information and other domains such as sensory and motor systems. This function introduces several constraints on the nature of these representations; irrespective of their content, they must be in a format that is readily accessible by both a range of linguistic and non-linguistic modalities of input and output, they must permit processing rapid enough to support on-line production and comprehension of meaningful speech at a rate of several words per second and to support rapid and appropriate motor responses to meaningful sensory stimuli in the environment, and they must enable flexibility of meaning across different contexts of use* (…)

In Krivochen (2011, 2012a, b, 2013) we stressed the issue of *format* as a *sine qua non* condition for *n*-objects to enter a dependency relation (for instance, via *concatenation*). At the level of semantic structuring, this seems particularly relevant, insofar as the result of the computations at C-I must be

[5] The following quotation is eloquent at this respect: "*According to Jackendoff, semantic structures (i.e. the semantic content in the mental lexicon) are simply a subset of conceptual structures which can be verbally expressed"* (Moss, et. al. 2007). However, *contra* Jackendoff, we think that the set of possible conceptual structures properly contains the set of *verbally* expressible structures, because of limitations on PF that are partly due to its essentially Markovian character.

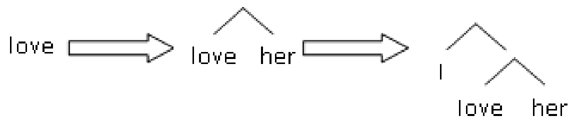
readable by more than a single system. The caveat with respect to the relation between lexical information and other systems is essential here: conceptual structure does not equate to lexical structure, rather, CS is ontologically and derivationally prior to lexical structure. The CS-lexicon interface is only one of the interfaces CS is conditioned by, and in turn, CS is only part of lexical information (which would arguably include categorial and distributional information in lexicalist models, both of which are absent in CS). A direct interface[6] between CS and "motor systems", if these are not directly connected to organs used for articulating speech, could result in an action rather than in a verbal externalization (which, incidentally, might eliminate the need for periodic finite-state cycles, since in this framework such cycles are phonologically motivated). Without entering a complex network of systems and interfaces like the one proposed by Jackendoff (2002, 2010) and his *parallel architecture*, where linking rules are indispensable (but quite anti-economical, if the model is to be implementable along the lines of Marr, 1983); or HPSG's networks (Green, 2011), to mention but two, it seems that a *multiple interface* approach to CS is, at the very least, hard to avoid (see also Uriagereka, 2008: 27-31).

The stance defended here is strongly aligned with *componential* models of meaning, which decompose complex meaning into more basic elements, in our case, roots (conceptual elements) and procedural / functional elements indicating the relevant system how to process the relations between roots, in line with Relevance Theory, see Sperber & Wilson, 1995; Wilson & Sperber, 2003; supported at the lexical level by psycholinguistic literature; but extending it beyond the concept to analyze how these conceptual and procedural are syntactically related (Uriagereka, 2008: Chapters 2 and 8; Hale and Keyser, 2002; Mateu Fontanals, 2005; among many others). This kind of conceptual structure is what, we propose, drives a linguistic derivation in a global sense, determining what is to be concatenated with what in order to convey the information of the CS in the less entropic possible way. The choice of elements to derive a sentence with, for instance (what is called an *array* or *numeration* in orthodox Minimalism; a subset of the whole lexicon in a particular natural language), is in our model determined by the CS and the information it conveys, insofar as the optimal scenario would seem to be that in which only the minimal number of elements *required* to convey CS in the least entropic way possible is selected from the Lexicon, regardless (at this point) of their materialization. While it would be at least odd to claim that a linguistic derivation starts with a phonological matrix, it seems to us to make sense to claim that it starts off with an *intention*, in the form of a complex CS.

However, a further variable is to be taken into account. If the structure is to be materialized (i.e., Spelled-Out), then the symbolic structures generated by *concatenation* must be not only materializable by virtue of their *format*, but also of their *content*. This means that Spell-Out cannot linearize certain phrase markers (a position argued for by Uriagereka, 2002a, 2012) because they fail to comply with a structural requirement (by virtue of being non-monotonic), but it is also possible that a terminal node cannot be provided with a phonological exponent because, in a certain natural language, there is no phonological exponent that corresponds to the information conveyed by a certain terminal or set of terminals. These situations are exemplified below (arrows signal successive derivational steps):

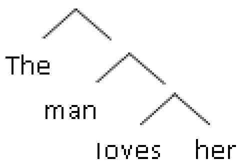
[6] An important note is in order: "direct" means that there is no third system mediating the connection, not that the interface is "transparent", as no two systems manipulate the same kind of information in the same format. Transfer, understood as information flow between systems, is always entropic. An important point a theory must address, in our opinion, is *how* entropic an interface is, and *to what extent* this entropy is countered by the organization of the system and its internal dynamics.

15) a.



b. Spell-Out input: [I [love [her]]]

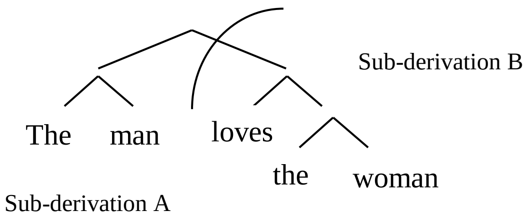
16) a.



b. Spell-Out input: *[The [man [loves [her]]]]

As we can see, the failure in phrase marker (16 a) is that [the man] is not treated as a constituent, either semantically or phonologically (e.g., in unmarked contexts, neither [the] nor [man] can be prominent by themselves, whereas the whole constituent could be prominent if the thematic position were relevant for some reason). For all syntax cares, being a component-neutral algorithm, it is a perfectly plausible phrase marker, but the semantic scope of the determiner over the noun is not captured (which is 'procedural' insofar as it indicates to the semantic interface how to interpret the semantic substance conveyed by the conceptual root categorized as a noun N, following Escandell & Leonetti, 2000). There is a hierarchy that is not captured via pure monotonic concatenation, whose limits, as Uriagereka (2012: 53) suggests, are those of a finite-state grammar (see also Karttunen & Beesley, 2005 for an overview of finite-state models of morpho-phonological properties). Conversely, if syntactic complexity increases, via non-monotonic concatenation (involving two non-terminals), Spell-Out must apply several times throughout a derivation (Uriagereka, 2002a), targeting each finite-state-compatible sub-derivation separately (see Toyoshima, 1997 for a similar idea, but employing different terminology):

17) a.



b. Spell-Out pattern: [[The man] [loves [the woman]]]

Since the global phrase marker goes beyond the limits of finite-state dependencies or Markov chains, Spell-Out applies to each finite-state-compatible sub-phrase marker (in this case, [the man] and [loves the woman], each of which has been assembled *separately* via monotonic concatenation in parallel workspaces, see Uriagereka, 2002a, 2012; Krivochen, 2015 a, b), thus chunking the derivation according to local PF requirements. The *parallel workspace* theory partly derives from the so-called "Finite State limit on phrase structure" (Uriagereka, 2012: 52, ff.), and, partly following –and expanding on- Idsardi & Raimy (in press), as a condition over symbolic representations, at a *local* level, phrase by phrase. The crucial idea is, as Uriagereka (2012: 88) puts it, is that '*the system is forced to undergo Spell-Out whenever phrases surpass the FS* [finite-state] *limit*'. This limit is reached when, for a syntactic object K in the form of a phrase marker, there is a symmetrical bifurcation of branches, as in (17 a) between [the man] and [loves the woman]. While this condition seems to play no role in the development of formal grammars, as we could easily imagine an L-grammar whose set of rules determined a symmetric bifurcation at each derivational point, the kind of grammars that seem to be significant for language display this kind of asymmetry, alternating *strong* and *weak* branches: relevantly, the kind of Fibonacci grammars studied by Uriagereka (1998: 192-193; 484, ff.); Shirley (2014), among others, are fundamentally asymmetric in their rule set[7]. Periodicity in terms of cycles is also related to this alternation, both from a PF (Uriagereka's MSO model) or an LF perspective (closer to our own).

Even at a smaller scale, let us say at the level of the word, morpho-phonological conditions can influence the derivational procedure, provided that we work with a system that allows dynamic access from the C-I and S-M systems to the working area. Single concept-level is the focus of attention of conceptual approaches like Moss et. al.'s (1997) Conceptual Structure Account, but here we will not make a concept-word correlation (as vaguely hinted at by Moss et. al., 2007): we focus rather on what constitutes a word, independently of the concept it instantiates. As has been argued by linguists for decades now (see Di Sciullo & Williams, 1987; Halle & Marantz, 1993; Hale & Keyser, 2002; Starke, 2011; Acedo-Matellán, 2010; Ackerman et. al. 2011, for a variety of views, within different frameworks, often not mutually compatible), a phonological word does not always correspond to a single terminal in a finite-state structure. Assuming that all terminals in the structure below convey some interpretable information (i.e., they are not "radically empty", cf. De Belder & van Craenembroeck, 2011), phonological insertion could target more than a single node. For example, the Spanish verbal morpheme [-ba] comprises Person (1^{st} or 3^{rd}), number (singular), Tense (past), Aspect (imperfective), and Modality (realis). Some languages, like Latin, allow even diathesis to be grouped with these nodes, such that so-called "passive voice" in the Infectum is expressed in a synthetic way, instead of English or Spanish analytical (periphrastic) mechanism (e.g., while Latin [amo] is 1^{st} person, singular, present, realis, active; [amor] comprises the same person, number, temporal, aspectual, and modal features, but *passive* voice). The procedure by which several terminals are materialized by a single morpho-phonological piece is called *fusion* in Distributed Morphology (see Halle & Marantz, 1993 for details).

Closely related to this, in past works we have proposed that, if the interfaces have access to the working space, then (*ceteris paribus*) they can condition at least part of the succession of concatenation operations performed there. It has been argued that terminals can be merged together, via *incorporation*, a sub-type of so-called *head-to-head movement*, in a traditional X-bar theoretical framework (see Baker, 1988 for extensive discussion; also Hale & Keyser, 2002 for discussion about the related operation *conflation*). In this operation, two terminals, α and β, are joined together to form a single complex terminal, known as an X^0_{max} (because it is neither a terminal in the strict sense; thus not an X^0, nor a fully-fledged phrase; thus not an X_{max} or XP), in Bare Phrase Structure terms (Chomsky, 1994). We have hypothesized, following much work in Distributed Morphology, that feature grouping in a terminal (where "features" could be interpreted in the sense of Taylor et. al., 2011, but including event-related features, like [manner], [motion], [direction], etc.), or, more generally, Spell-Out (if we consider that materialization can target more than a single terminal at a time, somehow on the line of Nanosyntax, see Starke, 2011), is locally conditioned by the possibility to actually insert a phonological exponent in that position, derived from the availability of a distributionally specified exponent. The principle we proposed determining the possibility of feature *fusion* is the following (adapted from Krivochen, 2012: 103):

[7] The relevant grammar is specified as follows:

0 → 1 (weak term)
1 → 0, 1 (strong term)

See Uriagereka (1998; 2014: 370) for a linguistic discussion, Saddy (2009) for a psycholinguistic discussion.

18) *Morpheme formation constraint*: syntactic mechanisms do not group features in a terminal node if there is no vocabulary item VI specific enough to be inserted in that node and Spell-Out the features present there.

For example, and following the traditional *verb-framed* vs. *satellite-framed* distinction (Talmy, 1991), Spanish (and other Romance languages) does not allow [manner] incorporation onto [motion], which we explain in our model by saying that there is no VI that can Spell-Out both features. However, it allows [direction] to incorporate onto [motion], whereas English (and other Germanic languages) behaves the other way around[8]. Examples like these abound in typological literature, and this particular phenomenon has been studied thoroughly by Mateu Fontanals (2002). This is, we think, a clear PF-constraint, acting at a local level. The tension between global semantic tendencies and more local morpho-phonological cycles is hopefully clear at this point. Let us reinforce our point with the introduction of a further view on cycles emerging from a fundamental tension: Zipf's view.

Zipf (1949: 35, ff.) explores the applicability of a (classical) harmonic motion equation (based on a Taylor series) for the '*group behavior of* (…) *a vocabulary of words in the stream of speech*' (1949: 37). In Zipf's terms, there is an *optimization* process between competing and opposite principles, in the case of a harmonic oscillator, relating size of the oscillator and frequency, '*in which the saturated harmonic equation will precisely reveal itself*' (1949: 38). Applied to language, Zipf formulates (admittedly informally) two principles, or rather Forces, what he calls 'Unification' and 'Diversification':

> <u>Force of Unification:</u> minimize the number of words (to 1), maximizing their frequency (to 100%) and associated meanings
>
> <u>Force of Diversification:</u> maximize the number of possible words, minimizing their frequency and associated meanings (tending towards a bi-univocal relation between words and meanings)

These 'forces' are usually identified with principles of economy centered in the speaker and hearer respectively (and have had long-lasting echoes, including Horn's 1988 Q and R Principles; and Uriagereka's 2012 treatment of the different perspectives within his CLASH model). Based on quantitative evidence from the Ulysses and the Iliad, Zipf claims that '*in the language of those two terms we may say that the vocabulary of a given stream of speech is constantly subject to the opposing Forces of Unification and Diversification*' (1949: 21), a perspective that is not alien to Harmonic grammars or bidirectional constraint interaction models (e.g., Jäger, 2002). His research on the relation between these opposing principles, and the optimal way to solve that tension lead him to say that

> *If we explicitly assume that the harmonic seriation $F \cdot Sn$* [expanded in the form of a Taylor series] *represents a fundamental principle that governs the number and frequency of usages of words in speech, then we can only conclude that a given speaker 'naturally' selects both the topics of his conversations and the words with which he verbalizes them in such a way that the resulting frequency-distribution of his continuing stream of speech will meet the exigencies of our equation, $F \cdot Sn$, without 'too little' or 'too much' talk. And this, in turn, means that inherent in the stream of speech is a dynamic unit which we may call a closure (or a cycle, or a*

[8] While it is true that the English lexicon does contain words like [enter] and [exit], which display both [motion] and [direction], they are actually Latin roots, not Germanic roots. Thus, both [enter] and [exit] are originally compounds of Latin [ire] ('go') plus a preposition indicating direction (in+ire; ex+ire).

> *rhythm) which might be defined roughly as the length of speech during which a particular group of verbal tools has completed its collective behavior once. What else this closure may signify we do not yet know* (…) (Zipf, 1949: 38-39)

Like Uriagereka, Zipf bases the resolution of the tension on linearization requirements (thus the frequent mention to the 'stream of speech'), interestingly, a decade before Tesnière's (1959) considerations about the different nature of semantics and morpho-phonology from a structuralist perspective. To the best of our knowledge, Zipf's work is the very first attempt to characterize a notion of cyclicity in natural language, and, what is more, to link it to a fundamental tension rooted on deeper properties of the system from which language emerges (even though their nature was unclear at the moment: where do Unification and Diversification come from?), as well as considerations of usage. It is to be highlighted, furthermore, that the notion of cycle Zipf assumes has strong cognitive correlates, since '*schizophrenic speech can almost be characterized by the absence of closure*' (1949: 39). Whether these considerations actually apply or not to specific pathological conditions (and see Zipf, 1949: Chapter 7 for more details) is not relevant for our purposes, insofar as we are setting our focus on the physical architecture that both *generates* such a tension between orthogonal components (the 'hiatus' or discontinuity, in the words of Uriagereka) and *solves* it locally (i.e., via local domains) in an optimal way. Interestingly, while we *can* in theory force the system to favor one interface over the other (or one force over the other), those are, according to Zipf (1949: 284, ff.) characteristics of pathological conditions (autism, schizophrenia) rather than the normal state of affairs in cognitive development. This is to be expected if a dynamical frustration is at the core of both cognition and its physical bases; and also if cognitive impairments impacting on language are *mapping* impairments (Kosta and Krivochen, 2015) which directly affect the available resources to solve the frustration between conceptual and morpho-phonological requirements. In past works we have hypothesized that

> (…) *each materialized linguistic stimulus is the best solution to the tension between what the speaker wants to say, that is, the information he wants to convey (information that is syntactically structured); and the means he has to externalize that meaning, which implies a structural flattening from hierarchy to linear dependencies.* (Kosta and Krivochen, 2015: 41)

The connections between our conception of the cycle, the notion of dynamical frustration, and Zipf's theory should be, at least, apparent. In the present work we have focused on the 'mentalist' aspect of the process and the role of cycles as emergent properties of the derivational dynamics we have described (in turn constrained by more basic physical principles), but a more complete picture of the implementational level of the theory should of course include considerations of language use in context, which, in the terms we have been presenting in these last paragraphs, amount to choosing the best candidate c_x (from a set of possible candidates C = {c_1, c_2, c_3, …, c_n}, *à la* Optimality Theory; see Prince and Smolensky, 2004) in a bidirectional system that solves the linearization-hierarchy conservation tension in an optimal way *for a speaker S, given finite resources* (working memory, time, lexical knowledge…) *in a communicative/situational context*. In such a dynamics, the orthogonal tension between principles Q and R should also be included, yielding a multidimensional dynamics, in which several phase spaces intersect to result in an externalized, linearized structure which conveys meaning *for a subject, to another subject, at a moment in time, and in a situational context.* This said, we will keep formalizing what happens 'within a speaker's head', although the importance of an extension of the theory along these lines cannot be overstated.

3. Conclusion: Unifying Frustrations

Throughout this paper, we have tried to keep the two frustrations we claim exist in the human mind separated for methodological reasons, since they are actually independent hypotheses. One could claim that the human mind displays tension between high- and low-level computational processes, without adhering to the theory that such a tension has any correlate to the semantics-phonology tension present in human language, and there is no logical inconsistency in that position. As we have hinted throughout the paper, also, we adhere to both hypotheses and, what is more, we claim that they can be unified in a single frustration, displaying variations of *scale* but not of nature: interactive, nonlinear processes are better suited to account for global semantic requirements, whereas they assign too rich a structure to externalized sound strings. The importance of bearing the notion of scale in mind cannot be exaggerated. The literature on quantum models of the mind, for instance, (e.g., Stapp, 2009; Penrose, 1997; Vitiello, 2001, among others) tends to make a generalization of uniformity about the quantum nature of mental processes, which makes their proposals vulnerable to Litt et al.'s (2006: 1-2) objection:

> *We argue* […] *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.*

Our proposal of *co-existence* of mixed processes (sometimes, even in the same phrase marker, as in 17, displaying two cascades) escapes this objection (as, for example, does Hameroff's 2007 mixed proposal, a reaction to Litt et al. 2006), which does not take into account the concept of *scale*. At an appropriate level of abstraction, it might very well be relevant how protein bonds configure structures that allow physiological processes at higher levels.

Provided that semantic/conceptual structures display dynamic, interactive properties; and that the linearization of such structures is plausibly conceptualized as cyclic Markovization (as put forth by Uriagereka, 2012, 2014; also Idsardi & Raimy, in press), the present proposal involving the unification of both *architectural* and *derivational* frustrations is, while not obvious, quite straightforward. Taking into account the architecture depicted in (13), it might very well be true that "*frustration mechanisms articulate representations*" (Uriagereka, 2012: 248), but we do not agree with Uriagereka that those representations fall *uniformly* within the Chomsky hierarchy. The explicit introduction of non-linear and interactive processes (the latter, for instance, to account for cyclicity effects, as in Figure I) goes beyond the limits of the Chomsky hierarchy, and we have argued in Krivochen (2015a) that the hierarchical taxonomy of formal languages must be thoroughly revised, if not altogether abandoned: while it is true that higher-level languages can express, or generate, everything lower-level languages can, and more; it is also true that most of the times the adoption of a uniform computational template entails the assignment of extra structure to the representation, an unjustified step, as Lasnik (2011) points out regarding the syntactic representation of iteration and the inadequacy of binary-branched X-bar theory for iterated structures (see also the criticism in Culicover & Jackendoff, 2005 of *structural uniformity* and their 'flat structure' alternative proposal).

3.1 *Squeezing n-dimensions through Spell-Out:*[9]

If we build up the dimensionality of syntactic representations by means of predicate-argument relations (as we have proposed in Krivochen, 2015a; based in part on Uriagereka, 2002b), linearization could proceed in a specular fashion: there are both empirical and theoretical arguments in favor of cyclic approaches to linearization, the gist of which (despite superficial differences) is that a derivation is chunked and linearized not as a whole, but proceeding chunk by chunk. Although there is relative consensus with respect to the necessity of a mechanism to implement cyclicity in computations, the specifics are far from agreed upon, including the size and identity of the relevant chunks (alternatives range from *v*Ps and CPs in Chomsky's *phase theory*, to finite-state sub-derivations in Uriagereka's MSO model; including Grohmann's 2003 quasi-cartographic *prolific domains*, Marantz's 2008 phases-within-words proposal, also Bresnan's 1971). In Section 1, we proposed our own version of what cycles might look like; now, we will attempt to provide a rigorous framework for their implementation. The relevant thesis from Krivochen (2015a) was that:

19) *If a predicate is to have scope over a referential (either sortal or eventive) variable, the number of the predicate's coordinates in the mental working area properly contain the number of the arguments' coordinates.*

'Arguments' are defined in this context as perturbations within the lexicon, understood as a field (see also Uriagereka, 2011 and Schöner, 2009 for a view of cognitive fields from Dynamical Field Theory (DFT)); whereas 'predicates' are the elements that, in a particular derivation, make those perturbations (formalized as wavefunctions) interfere, thus generating specific patterns depending on the elements that are related. For instance, let us see what a predication relation over a sortal entity defined in two dimensions would look like:

20) a) N = (x, y)

b) A = (x', y', z)

c) A(N) = A × B = $(x, x')\ (y, y')\ (\emptyset, z)$[10]

And so on. For instance, if a category X (X ≠ A) takes (A(N)) as an argument, it will have to be defined in (x'', y'', z'', w). As the reader may have noticed, the predication relation involving coordinates is expressed as the cross product of the coordinate sets involved, a move that makes sense if we are dealing with fields: in Dirac notation, operations over vector fields are cross product vectors, such that the inner product of a *bra* and a *ket* vectors $\langle\alpha \mid \beta\rangle$ is $\langle\alpha \mid \times \mid \beta\rangle$. If this approach turns out to be on the right track, many essential properties of the generative (structure-building) operation will follow from the very architecture of the cognitive system. This will allow us to define very precisely the scope of multiple modifiers over an argument, as in iteration (e.g., if *n* As modify an N, but none of them has scope over another, then they are all defined by the same *number* of coordinates in the cognitive workspace, the respective values for the individual components varying as these values represent the position the relevant lexical vector points to within the field, which translates to a finite-state representation, of the kind we have proposed for iteration and some instances of coordination in Krivochen, 2015a), thus effectively tackling the problem of 'too much structure' noticed by Lasnik (2011). At this point in the development of the proposal, it is not really crucial whether N has one,

[9] This section is based on work in progress, and makes reference to issues that we have discussed at length in Krivochen (in preparation). The interested reader is welcome to contact us for details.

[10] The result could have also been expressed in matrix notation. The reader interested in linear algebra is welcome to try the exercise.

two, or n coordinates, just that if N = (n) (i.e., N is defined in n dimensions), the predicate that takes N as its argument will be defined as (n+1). In other words, *for n dimensions of α, f(α) = n+1* (for a different perspective, see Uriagereka, 2002b: 296, who proposes that spatial dimensions can 'warp' in order for the manifold representation to get to a higher-level dimensionality; similar arguments are used in string theory, with dimensions stretching and folding in specific topological configurations). It is not clear that the relevant space has 'warped', but, rather, that n+1 properly contains the relevant object described within n, such that a predication relation can be established. The 'warping' mechanism does not help with linearization either, in the terms we are working with here (but see Uriagereka, 2011: 49 for a proposal within the 'kite model', quite reminiscent of so-called 'cusp catastrophes' in catastrophe theory, see Zeeman, 1977), as it is not clear how the warping is actually implemented in a dynamical system. Moreover, recent developments of Uriagereka's theory are based on the assumption of the ultrametricity of the syntactic space (see also Roberts, 2015 for a more mathematically oriented proposal focused on phrase structure), which is to be relativized if syntactic relations impose a metric over the elements involved, making them be 'closer' to one another than with to other, derivation-external objects of the lexicon. Here, we will attempt to adapt an idea from Schöner (2009) about the properties of unstable n-dimensional dynamical systems, the so-called 'Center Manifold Theorem' (see also Guckenheimer and Holmes, 1997, Theorem 3.2.1) in order to implement the notion of *cycle* derived from a *dynamical frustration*:

> *When an attractor of a high-dimensional dynamical system* […] *becomes unstable there is typically one direction in the high-dimensional space along which the restoring forces begin to fade; while in other directions, the stabilization mechanism remains strong. The one direction along which forces fade spans the center manifold.* (Schöner, 2009: 9)

Haragus and Iooss (2011: 34), within a more rigorous mathematical framework, also claim that the center manifold theorem reduces the dimensionality of a topological object, as center manifolds arise near critical points in a dynamical system, which will be crucial for our definition of cycle insofar as these critical points actually *delimit* the domains for the application of operations and the establishment of dependencies between objects. Zeeman (1977) narrows down the scope of the mathematical discussion to the interplay between neural structure and cognitive processes, focusing on discontinuities, which he claims is a major characteristic of the systems of differential equations that model neurocognitive dynamics. Crucially, while some properties of high-dimensional complex systems might be difficult to solve with ordinary differential equations, solutions in a center manifold can always be described by a finite-dimensional system of such equations, which is indeed an advantage from both theoretical and empirical perspectives.

If a derivation can be described by a function relating aspects of structure building (like dimensionality of the manifold) and informativity (like entropy), the function corresponding to a derivation is *discontinuous* due to the appearance of cycles which represent a temporary solution to the dynamical frustration between semantic and materialization requirements as we have argued above. The critical value at each cycle, call it p, delimits the cycle, on the one hand, and defines a point of non-differentiation on the other. Given a discontinuous function at p; $\frac{dy}{dx}(p)$ is undefined and thus the system loses stability (causing an attractor to disappear or fade, insofar as $\Delta x \approx 0$). Crucially, we will propose here that the critical value p is the point where a cycle ends, thus motivating the assumption that a cycle is spelled-out (i.e., linearized and materialized) once completed, there being a discontinuity (a 'hiatus' in Uriagereka's terms, a 'catastrophe', in Zeeman's terms) in the dynamics of the system, to be solved in a lower dimensionality.

The direction (i.e., the axis) in the phase space along which the unstable attractor fades define the low(er) dimensional *center manifold*, a number of orbits whose behavior around the equilibrium point is not determined by either attractors or repellors (points, or set thereof, towards which or away from which the system tends, regardless of its initial state (Schöner, 2009 for a DTF perspective; Aguayo et al. 2004 for a mathematical physics perspective and formalization), is the direction along which the system evolves slower (as restoring forces are weaker), which means the rate of change $\frac{\Delta y}{\Delta x}$ is smaller, for any dependent variable we want to consider: in the case of the system we have been presenting here, we could consider *y* to be *entropy* or *dimensionality* (of the structural manifold).

For this argument to work, we first have to make sure we are in fact dealing with a (differentiable) manifold: are objects derived by means of syntactic mechanisms manifolds? Consider the argument in favor of the increasing dimensionality of derivations by means of predication relations, such that *for n dimensions of an object α, f(α)* [a predication relation taking α as an argument] = *n+1*. This system predicts each structural 'generation' will have a higher dimensionality than the previous one(s), provided (19) holds. A derivation would expand in a multidimensional space in real time as it is built until reaching a critical value, then, an attractor would fade, which determines the direction in which the system 'squeezes' into a lesser-dimensional space. How to determine this point? We said that it is the point at which the function is discontinuous, but that alone would make a circular argument. The question is: why should the function defining the global behavior of the system over time be discontinuous (and, more specifically, discontinuous in the way we suggest here)? The answer, in our opinion, lies in the concept of *dynamical frustration*, which makes its glorious reappearance here.

Recall we have followed Uriagereka (2011, 2012, 2014) in claiming that there is a fundamental tension in the architecture of language as a cognitive system between conceptual and morphophonological trends, which is one of the main features of dynamical nonlinear systems (Binder, 2008). More specifically, we have argued in Section 1 in favor or a *scale frustration*, in which global trends favor conceptual conservation principles (see also Lasnik and Uriagereka, 2005: 53; Krivochen, 2011: 53) from the Conceptual Structure[11], which is non-linguistic, through its linguistic instantiation all the way towards the so-called 'Logical Form' in the architecture sketched in (13) (we take LF here to be more in relation to Relevance-Theoretic LF than to the narrowly first-order logic representation that is customary in Generative Linguistics, see Hornstein, 1995 for a review of LF within Generativism). Local trends, in contrast, favor cyclic aspects of computation (including cyclical rule application), targeting minimal domains. We also said that we reject the traditional generative Y-model in which the syntactic component operates over elements and *sends* information to the interfaces, since (a) on both empirical and conceptual grounds, we reject the autonomy of syntax and its alleged 'specificity' as there is structure, and therefore syntax, throughout cognition; and (b) we argue for a dynamical model in which the interfaces (the sound and meaning systems, for concreteness) *access* the derivational workspace and *take* whatever they can minimally read (instead of just passively *receiving* whatever the syntactic component sends).

An immediate consequence of this proposal is that cycles are not intrinsic to the so-called computational system for human language $C_{(HL)}$ (*contra* Chomsky, 2008 and much related work), nor

[11] In turn, the conceptual structure is the result of syntactic operations applying within the conceptual space, which we have characterized as an *n*-dimensional ultrametric phase space. For ease of visualization, and using Douglas Saddy's expression, we are in presence of a 'conceptual sabayon' whose bubbles (i.e., empty space) configure the space of what is not thinkable.

are they based on purely 'formal features' (cf. Gallego, 2010: 51; Chomsky, 2001, 2008) or any other intra-theoretical stipulation, but arise from the frustrated nature of the system: cycles are the emergent result of the self-regulation dynamics of the nonlinear complex linguistic architecture we assume. Morphophonologically, cycles are materializable chunks of structure, sets of solutions within the phonotactic space at the segmental level and within the ultrametric lexicon at the word level (although the concept of 'word' is far from clear, we include here both free and bound morphemes, and will not get into further details with respect to the lexico-morphophonological component. Despite global conservation trends, there are also semantic cycles, which we have referred to in past works as 'informational domains' (Krivochen and Kosta, 2013: 162-163; Krivochen, 2011: 50; 2012: 30-31). Those informational domains are built incrementally, proceeding step by step in a process that, as time (t) goes by, increases dimensionality (D) and makes entropy (h) decrease (since there are more clues for the interpretative procedures with respect to where in the phase space to look for relevant attractors). Let us now introduce some formalization regarding the cumulative structure building system.

If we depart from the simplest case, in which the first element to enter the derivational space is defined by means of a single dimension d_x defined on the x axis (a case we will consider only for expository purposes), the derivation will proceed increasing the dimensionality of the structure being built arithmetically. Assuming a simple monotonic structure of predication, that process could be graphed as follows:

21) $d_x \rightarrow d_{x,y}(d_x) \rightarrow d_{x,y,z}(d_{x,y}(d_x)))\ldots$

That is, in order to characterize the third derivational stage exhaustively, we need to resort to a description of the field's wavefunction ψ_D in three dimensions, which using Dirac notation would be a 1 × 3 matrix:

22) $|\psi_D\rangle = \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix}$

This means that each derivational step has a vector associated to it (in Dirac notation, which we have used in (22), it would be a *ket* vector, a column of dimensional components expressed as real values), with as many subindexes as minimally required in order to describe exhaustively the state of the system at that particular derivational point[12, 13]. If there were no cycles, dimensionality would quickly grow beyond P / NP computability (even in an unrealistically simple case as (21)); *a fortiori* since monotonicity is not quite the norm in linguistic derivations: consider not only the case of subjects (which can be monotonically assembled structures themselves, as in (17)), but also coordination, both asyndetic and overt (Krivochen, 2015a); adjunction (Uriagereka, 2002a;

[12] We have already hinted at the claim that any strict quantization of time is artificial due to the essentially continuous nature of the processes we are modelling, but we feel the need to stress it here: we take derivations to proceed incrementally over quantized time (as we said above, t adopts integer values) just for purposes of exposition. We do not forget, however (and nor should the reader), that we are dealing with continuous, interactive processes. We think this work is only a first step towards a more articulate theory of linguistic derivations in dynamical field terms, which should eventually deal with the continuous nature of time (and other nuances we have overlooked here) due to the finiteness of both the time available to complete this work and our cognitive capacities.

[13] We are not taking into account *bra* vectors, since we are not dealing with complex conjugates: take the relevant *bra* vector to be of the form $1 + i0$ (no imaginary part), trivially, the complex conjugate is $1 - i0 = 1$.

Krivochen, 2015a); path-of-motion and resultative constructions (Krivochen, 2012: 71, ff; Krivochen and Kosta, 2013: 115-119), parasitic gaps, and multiple gap constructions (Krivochen, 2015b), among other phenomena involving more than a single derivational space[14]. In these cases, we are relating derivational cascades (using Uriagereka's terms) each with its own derivational history, thus, their own associated *ket* vector. An interesting constraint that arises from our proposal is that, if non-monotonic relations are expressible in terms of operations over the respective *ket* vectors, we can only non-monotonically relate objects with the same number of dimensions; otherwise, we cannot operate over those vectors. However, we cannot stipulate this *a priori*, since syntactic relations involve a variable number of elements defined by *n* dimensions, and as we are dealing with an open, interaction-based, nonlinear dynamical system, there are essential properties of the system that prevent us from making deterministic statements like that (among other properties, we count hypersensitivity to initial conditions and emergence). What are we to do, then? One approach, is to say that it is precisely this incompatibility that weakens a dimensional attractor in the higher-dimensional object, thus triggering its 'dimensional squeezing' through the weakened axis. On this approach, the center manifold theorem would follow from the architecture of the system in a natural way, yielding cyclicity as a by-product of the system's dynamics (a highly desirable by-product, both empirically and theoretically, as we have argued). Naturally, the points in which there is a dimensionality incompatibility are non-differentiable, because the 'cumulative' nature of structure-building operations is halted due to the rearrangement of the system around the stable center manifold: as we said before, if p is the (derivational) point at which the dimensionality tension is to be solved, and the solution involves a reorganization of the system around different attractors, $\frac{dh}{dt}$ is not defined at p. Interestingly, from the perspective of multi-dimensional field theory applied to neural dynamics, Schneegans (2014: 2) claims that multidimensional fields require an '*interaction kernel of the same dimensionality*', even though the mathematical possibilities for higher-dimensionality are 'straightforward'. The constraints are due not to limitations in measuring capacity or theoretical shortcomings, but rather to essential features of interactions between multidimensional physical systems.

A direct consequence is that, as in Uriagereka's model, non finite-state units configure derivational cascades, but crucially, not because of their incompatibility with the LCA (since we find the function that relates command between terminals and linear order stipulative, as well as the limitations it imposes over phrase structure yielding an inevitably uniform system), but because the establishment of a dependency between monotonically derived units in parallel workspaces implies a dimensionality increase that acts as a critical point for the computational limitations of the system, triggering the reorganization of the system around the stable center manifold.

In terms of probability measurement (see, e.g., Feynmann, 2013: Chapter 5), the proposal above also makes sense: consider the *bra-ket* vectors in (23)

23) $\langle \alpha \mid \beta \rangle$

The physical interpretation of (23) in terms of measurement (which is relevant to our purposes) is relatively simple, and to exemplify it we will consider the case of electron spin. Assume we prepare an electron in spin β and measure it with a piece of equipment that detects spin α. The probability amplitude is given by the inner product of α and β; the actual probability, by $|\langle\alpha|\beta\rangle|^2$, that

[14] Certain operations in TAGs (Joshi, 1985) also fall within this category, at least formally: it is not obvious how the TAG mechanism could be implemented in a system like the one we propose here, since they are limited to PDA+ computation as we saw above.

is, multiply the probability amplitude by its complex conjugate (which is trivial if we are dealing with real numbers). Let us consider a relatively simple case, up and down spin; where β is the spin in which we prepare an electron (say, Up) and α is the state our detector measures (in this case, Down). Up is defined by the coordinates (0, 1)[15] whereas Down is defined by the coordinates (1, 0). Both *bra* and *ket* vectors are defined as matrices, such that:

24) a. $\langle \alpha | = (0, 1)$

b. $|\beta\rangle = \binom{1}{0}$

The inner product of the *bra-ket* vectors, as we said, gives us the probability amplitude, so let us perform the relevant operations which give us a scalar:

25) $\langle \alpha \mid \beta \rangle = 0 \times 1 + 1 \times 0 = 0$

This means, as we would expect, that the possibility of measuring a spin down electron with a spin up detector is 0 (in order to get the probability, we have to square 0, which of course equals 0). In more general terms, we have (26) (see Feynman, 2013: Vol. 3, Chapter 5, equation 5.25; Chapter 16, equation 16.3):

26) $\langle \alpha_i \mid \beta_j \rangle = 1$ *iff* $i = j$

0 *iff* $i \neq j$

Translated to physical terms, this means that if we prepare an electron with spin *i*, and measure it with a detector configured in spin *j*, the corresponding eigenvalue (i.e., the result of the experiment) will be 1 (i.e., complete certainty the electron will be detected) only if the spin configuration of the electron and that of the detector are equal (e.g., both Up, both Down). Otherwise (and limiting ourselves to a two-outcome system[16]), the probability of detecting an electron is 0. Crucially, in more complex systems, where the *bra-ket* is just one component, we *have* to assume $i = j$, otherwise, the whole term just cancels. For instance, assume we have a wavefunction $|\psi\rangle$, and we want to know the prospects of preparing the system in state ψ and measuring it in state x_j. That is expressed as follows (see, e.g., Littlejohn, 2007; Feynman, 2013: Vol. 3, Chapter 5, equation 5.24; also Chapter 8, equation 8.1):

27) $\langle x_j | \, \psi\rangle = \langle x_j | \sum_i a_i \, |x_i\rangle = \sum_i a_i \, \langle x_j | x_i\rangle$ [17, 18]

If x_i and x_j are orthogonal to each other, then we have to assume $i = j$, otherwise, the result of the inner product will be 0, thus cancelling the whole term (since $\sum_i a_i \, 0 = 0$). If the *bra* and *ket* vectors indicate measured and prepared states of a system, respectively, we cannot have $i > j$ or *vice*

[15] Meaning: '*x* components down, *y* components up'

[16] The same procedure works for photon polarization. Assume we prepare a photon with a polarization *m* which is at an angle θ with respect to the measure *n*. The probability is $cos^2 \frac{\theta}{2}$. So, for instance, if *m* and *n* are at an angle θ = 0°, $cos^2 0° = 1$. Notice that in this case, $m = n$, which satisfies the equality in (110).

[17] Notice that we can take $\sum_i a_i$ out of the *bra-ket* notation because they are numbers –scalars-, not matrices or vectors.

[18] Naturally, if the changes over *x* are 'infinitesimal' in Leibniz's sense, we can always replace ∑ with ∫ (adding *dx*, of course), in which case *x* would be *almost* continuous by a factor Δ*x*.

versa. If, as we proposed, each dimension is a component in the vector, we can only operate with vectors that have the same number of components, or fix the missing components in one of the vectors to 0 (that is, say 'the system does not have a value for component *n*'). Such operation would be trivial: imagine we have the inner product of α and β, such that α is defined in 3 dimensions and β is defined in 2. The inner product would be:

28) $\langle \alpha \mid \beta \rangle = \alpha_x \times \beta_x + \alpha_y \times \beta_y + \alpha_z \times 0$

This would just decrease the dimensionality of the operation by 1, since the *z* component (i.e., the value of the *z* coordinate) in α would be neutralized. In our opinion, this is another way of saying that multidimensional field interaction '*requires interaction kernel of the same dimensionality*' (Schneegas, 2014: 2), but in a slightly different manner: if the 'same dimensionality' requirement is not met, the system forces a dimensionality reduction, which we have expressed in topological terms (via the center manifold theorem) and now in purely physical terms, via *bra-ket* notation. Now, is this dimensionality reduction relevant? We think so, as either we are forcing the rise of a cycle, or we are simply losing all the information in the neutralized dimension. If the dimensional incompatibility between a number *n* of *ket* vectors did not force the fading of an attractor and we just kept computing as if nothing had happened, we would be violating the *conservation principle* that globally rules derivations, by losing information (instead of squeezing that information through a lower dimensional complex plane).

A similar position with respect to the relation between cognitive processes and their externalization in terms of action or language is that of Spivey (2007), which we have introduced above. According to Spivey,

> *Action and communication* […] *necessarily overidealize and exaggerate the discreteness of people's internal representations, typically settling on the closest response category to what was originally intended.* (2007: 168)

In terms of externalization, then, there is an essential discontinuity -imposed by means of chunking and Spell-Out- over cognitive representations and processes, which are essentially continuous over time. This view is also in consonance with DFT, although the latter provides tools for describing the state of the field at an arbitrary discrete point or set thereof by means of quantizing *t* (see the differential equations in Schöner, 2009, for instance), which are most useful insofar as we keep in mind that such quantization is artificial. Interestingly, Spivey's quotation leaves open the interpretation that 'what was originally intended' is not a point, but a phase space on its own, and the 'closest response category', an attractor within that space. Notice that if the intention was a point (or even a vector, or a tensor) rather than a set of possible solutions that satisfy the problem to different extents (in a core-periphery dynamics), then there would be a single possible externalization *per* internal representation, that whose coordinates (the corresponding *ket* vector) coincide exactly with the representation's. In our opinion, such a deterministic model is too strong and constrained, and it is not clear how a linguistic representation (or any other cognitive representation) could encode the required specificity in terms of attractor localization. Our model is quite more flexible: even if a representation whose interpretation is to be related to an intention (e.g., for the purposes of inference extraction, where the hypotheses we make about the other person's intentions when producing a certain stimulus are crucial), there is an error margin which allows for the 'closest response category' to be located and which, for all practical purposes, might be very well equivalent to the exact point. This approach, we think, makes the process of finding an attractor not only 'easier', but also computationally less costly, which is a desirable result.

If the picture we have sketched above about variable entropy is correct, then if time expressed in positive integers (t = 1, 2, 3…) correspond to derivational cycles, then, $\lim_{t \to \mathbb{N}} h(t) = 0$. The function is limited by two asymptotes, 1 and 0 (complete disorder and complete informativity, respectively), and it oscillates between attractors at (t, 1) and (t, 0), where t is a positive integer. Since our attractor points belong to an asymptote line, Spencer et al.'s (2009: 109) proposal that the system goes *into* stable attractors, having enough flexibility to 'escape' the stable attractor and change the dynamics does not seem to be feasible under our assumptions, as no value of the function can ever reach a value within the asymptote line. Recall, however, that for expository purposes only we are taking t to have values that are integers, corresponding to derivational steps that are quantized in an equally artificial manner, to avoid an internal contradiction. The focus on *process* rather than *state* is a basic feature of interactive-based computation, and we would like to maintain it as far as we can, since we do think it has theoretical and empirical benefits. This is particularly case from the perspective that taking the first step towards an explanatory model of brain dynamics based on the notion of *tension* (i.e., *dynamical frustration*), which in turn is pervasive in (other) physical systems—just as *cycles are*—is too promising to be left unexplored.

It must be noted that, unlike Epstein et. al. (2010), we do not claim that there is any 'design' problem, such that the computational system is 'perfectly designed' to fulfill C-I interface conditions primarily, and S-M conditions only secondarily. That view, which derives from Chomsky's (2000) Strong Minimalist Thesis, entails, we think, a teleological conception of the computational system $C_{(HL)}$ and its interfaces. In evolutionary terms, it is not possible to say that the computational system evolved *in order to* fulfill C-I conditions; rather, evolution brought together two systems (S-M and C-I) and their intersection (think of it set-theoretically, for more clarity) has some specific properties (Conceptual Structures). In purely linguistic terms, Epstein et. al.'s argumentation revolves around the notion of *feature interpretability* and the role of Transfer in providing the interfaces with fully readable objects. By eliminating the role of features of the kind [*value* – Dimension] here and in previous works (Krivochen, 2011, 2012; Krivochen & Kosta, 2013), we reformulate the role of Transfer as conceived in the orthodox Minimalist Program (see (5)) as the interfaces *taking* rather than the syntactic component (which is not the center of the theory either) *taking*. Transfer timing is thus not determined by feature-valuation issues or stipulatively defined cycles (e.g., phase heads), but strictly by the interfaces' peering into the syntactic workspace and each one independently takling the minimal object it can assign a representation in terms of its own substantive elements (e.g., phonemes). Transfer is thus not an operation of 'feature stripping', as in many versions of Minimalism, but a set of multiple periodic step-by-step evaluation (*Analyze*) and 'co-optation' operations.

The program we have presented in this paper is, we hope, clear: natural language displays global semantic tendencies: syntactic derivations are semantically driven by the need to conserve the information present in the CS throughout the derivational path towards LF. At a more local level, however, morpho-phonological requirements determine the existence of cycles, which are also a characteristic of chaotic systems, as we saw in Figure I. At each derivational point, there is a tension between global semantic conservation tendencies and periodic morpho-phonological transfer points. In other words, there is an unlimited capacity for discrete object manipulation and combination in n dimensions, but possibilities for (phonologically) materializing the symbolic structures built this way that are limited by, on the one hand, the finiteness of the language's lexicon at any time in history and at any derivational point and, on the other, by the finite-state computational power of the phonological channel—a tension never to be solved in *equilibrium* (in Binder's 2008 terms) as it keeps a derivation going. Future research will determine if this tension is actually resolved by the system in the most economical way possible given global and local requirements so that the result thus is an optimal externalized expression of propositional content. This approach to cognition has enormous impact for theses of optimality of the design of mental faculties (Chomsky, 2005; Brody, 2003, among many others), since the optimal resolution of the frustration at each derivational point, which would yield an optimal output, does not necessarily entail "perfection" in the design of the mental faculty (or faculties) that produces that output. The fact that a dynamic system can, as we think language does, solve its inherent PF-LF frustration in an optimal way, does not mean that the system is, in any relevant way, "perfect". While the notion of optimality can be defined unambiguously, design perfection is relative to a system, a function, and an evolutionary trait (see Kinsella, 2009 for discussion). Only the belief that evolution (e.g., a putative mutation that yielded Merge) is driven by "virtual conceptual necessity" (an undefined and vague notion) can lead to the claim that optimality in the output implies perfection in the generative procedure and thus some kind of teleological "always-for-the-best" evolutionary process. The result of the resolution of the PF-LF frustration we have been analyzing, is "optimal" *if, and only if*, to cite Jäger & Blutner, 2000: 21, given a pairing of form and meaning (π, λ), where π = Phon and λ = Sem,

29) $(\pi, \lambda) \in$ GEN [the set of formulae that a given algorithm, like *concatenation* can generate]
There is no optimal (π', λ) such that $(\pi', \lambda) <$ [is ranked higher than] (π, λ), and
There is no optimal (π, λ') such that $(\pi, \lambda') < (\pi, \lambda)$

While this definition can be (and has been indeed) challenged, it has been developed extensively within frameworks assuming a system of constraint weight and interaction, like Optimality Theory. No such explicit development has been provided in favor of the "perfection argument" in mental faculties, a thesis that is, in our opinion, not even well formulated, as "perfect" is often confused with "optimal" (the former being absolute, the latter being relative to a certain measure, like interface conditions). As prospects for future research, we propose to redefine optimality taking into account the notion of *frustration*, and how the essential tensions between systems are solved; and to revisit the notion of perfection of a system based on the optimality of the output. We think that a mixed model of mental computations like the one proposed here is a good tool for the job, although, needless to say, not the only possible one.

4. *Bibliography*

Acedo-Matellán, V. (2010) *Argument structure and the syntax-morphology interface. A case study in Latin and other languages*. PhD Thesis. Departament de Lingüística General, Facultat de Filologia. Universitat de Barcelona

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.

Aguayo, J., M. Saavedra, and M. Wallace (2004) Attractor and Repeller Points for a Several-Variable Analytic Dynamical System in a Non-Archimedean Setting. *Theoretical and Mathematical Physics* 140(2). 1175-1181.

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

Bak, P. & M. Paczuski (1995) Complexity, contingency, and criticality. *PNAS* 92. 6689-6696.

Baker, M. (1988) *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

Belk, Z. & A. Neeleman (2013) AP-Adjacency as a Precedence Constraint. Ms. http://ling.auf.net/lingbuzz/001906

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

Boeckx, C. (2008) *Understanding Minimalist Syntax*. Oxford: Blackwell.

Brody, M. (2003) *Towards an Elegant Syntax*. London: Routledge.

Caramazza, A., & B. Mahon (2006) The organization of conceptual knowledge in the brain: The future's past and some future directions. *Cognitive Neuropsychology*, 23(1), 13-38.

Carruthers, P. (2006) *The Architecture of the Mind*. Oxford: Clarendon Press.

Chomsky, N. (1994) Bare phrase structure. *MIT Occasional Papers in Linguistics 5*.

(1995) *The Minimalist Program*. Cambridge, Mass.: MIT Press.

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

(2000) Minimalist Inquiries: The Framework. In R. Martin, D. Michaels and J. Uriagereka (eds) *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, Cambridge, MA: MIT Press, 89–155.

(2005) Three factors in language design. *Linguistic Inquiry* 36, 1-22.

(2008) On Phases. In *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*. R. Freidin, M. L. Zubizarrieta & C. P. Otero (eds.). MIT Press. 133-166.

(2013) Problems of Projection. *Lingua* 130. 33-49.

Citko, B. (2011) Multidominance. In Boeckx, Cedric (ed.) *The Oxford Handbook of Linguistic Minimalism*. Oxford: OUP. DOI:10.1093/oxfordhb/9780199549368.013.0006

Corning, P. A. (2002) The Re-Emergence of 'Emergence': A Venerable Concept in Search of a Theory, *Complexity* **7** (6): 18–30.

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

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

Dehaene, S. (1997) *The Number Sense: How the Mind Creates Mathematics*. Oxford: OUP.

Den Dikken, M. (2014) On feature interpretability and inheritance. In Kosta, P. et. al. (eds.) *Minimalism and Beyond: Radicalizing the Interfaces*. Amsterdam, John Benjamins. 37-55.

Di Sciullo, A-M & E. Williams (1987) *On the Definition of Word*. Cambridge, Mass.: MIT Press.

Epstein, S., H. Kitahara & T. D. Seely (2010) Uninterpretable Features: What are they and what do they do? In Putnam, M. (ed.) *Exploring Crash-Proof Grammars*. Amsterdam, John Benjamins. 125-142.

Feynman, R. P. (2006) [1985]. *QED: The Strange Theory of Light and Matter*. Princeton University Press.

(2011) [1963] *The Feynman lectures on Physics*. 3 Vols. Available online at http://www.feynmanlectures.caltech.edu/III_toc.html [Retrieved on 27/01/2015]

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

Fodor, J. D. (2013) Pronouncing and comprehending center-embedded sentences. In Sanz, M., I. Laka & M. Tanenhaus (eds.) *Language Down the Garden Path: The Cognitive and Biological Basis for Linguistic Structures.* Oxford: Oxford University Press.

Guckenheimer, J. and P. Holmes (1997) *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Berlin, New York: Springer-Verlag.

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.

Goldstein, J. (1999) Emergence as a Construct: History and Issues. In *Emergence: Complexity and Organization* 1 (1): 49–72.

Hale, K. & S. J. Keyser (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.

Haragus, M., and G. Iooss (2011) *Local Bifurcations, Center Manifolds, and Normal Forms in Infinite-Dimensional Dynamical Systems*. London: Springer.

Heck, F. & G. Müller (2007) Extremely Local Optimization. In: E. Brainbridge & B. Agbayani, eds., *Proceedings of the 26th WECOL. California State University, Fresno*. 170–183.

Hofstadter, D. R. (1979), *Gödel, Escher, Bach: An Eternal Golden Braid*. Basic Books.

Hornstein, Norbert (1995) *Logical Form: From GB to Minimalism*. New York: Wiley and Sons.

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

Jackendoff, R (1983) *Semantics and cognition*, Cambridge, Mass.: MIT Press.

(1989) What is a concept, that a person can grasp it? *Mind and Language*, 4**,** 68-102.

(2002) *Foundations of Language. Mind, Meaning, Grammar, Evolution*. Oxford: OUP.

(2010) *Meaning and the Lexicon: The Parallel Architecture 1975-2010*. Oxford: OUP.

Jäger, G. & R. Blutner (2000) Against Lexical Decomposition in Syntax. In Jäger, G. & R. Blutner (eds.) *Studies in Optimality Theory*. Potsdam Linguistic Investigations 8.

Joshi, A. K. (1985) Tree adjoining grammars: How much context-sensitivity is required to provide reasonable structural descriptions? In David Dowty, Lauri Kartunnen, and Arnold Zwicky (eds.) *Natural Language Parsing*. Cambridge, Mass.: CUP. 206-250.

Kartunnen, L. & K. Beesley (2005) 25 years of finite state morphology. In Carlson, L. & K. Ylijyrä (eds.) *A Festschrift for Kimmo Koskenniemi.* CSLI Publications. 1-13.

Kinsella, A. (2009) *Language Evolution and Syntactic Theory*. Cambridge: CUP.

Kitahara, H. (1997) *Elementary Operations and Optimal Derivations*. Cambridge, Mass.: MIT Press.

Kosta, Peter and Diego Krivochen (2014) Flavors of Movement: Revisiting the A-A' distinction. In Kosta, Peter, Steven L. Franks, Teodora Radeva-Bork and Lilia Schürcks (eds.), *Minimalism and Beyond: Radicalizing the interfaces*. Amsterdam: John Benjamins. 236-266.

Krivochen, D. (2011) An Introduction to Radical Minimalism I: on Merge and Agree. *IBERIA*, 3(2). 20-62.

(2012a) *The Syntax and Semantics of the Nominal Construction*. Frankfurt am Main: Peter Lang.

(2012b) Towards a Geometrical Syntax. Ms.

(2013) Language, Complexity, and Entropy. Ms. Under review.

(2015 a) Types vs. Tokens: Displacement Revisited. *Studia Linguistica* [Issue, Number, Pages]

(2015 b) On Phrase Structure building and Labeling algorithms: towards a non-uniform theory of syntactic structures. *The Linguistic Review* 32(3). [Pages]

Krivochen, Diego and Katarzyna Mathiasen (2012) *Numerals, Numbers and Cognition: Towards a Localist Theory of the Mind*. Poster presented at *Cognitive Modules and Interfaces*, international workshop at SISSA, Trieste, Italy. http://ling.auf.net/lingbuzz/001653

Lasnik, H. (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. and M. Saito (1984) On the nature of proper government *Linguistic Inquiry* 15: 235-289

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.

Lasnik, H., J. Uriagereka & C. Boeckx (2005) *A Course in Minimalist Syntax*. London: Blackwell.

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

Litthejohn, R. (2007) The Mathematical Formalism of Quantum Mechanics. Ms. Berkeley University. Retrieved from http://bohr.physics.berkeley.edu/classes/221/0708/notes/hilbert.pdf [27/01/2015]

Marr, David (1982) *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman.

Martin, R. and J. Uriagereka (2014) Chains in Minimalism. In Kosta, Peter, Steven L. Franks, Teodora Radeva-Bork and Lilia Schürcks (eds.), *Minimalism and Beyond: Radicalizing the interfaces*. Amsterdam: John Benjamins. 169-194.

Mateu Fontanals, J. (2002) *Argument Structure. Relational Construal at the Syntax-Semantics Interface*. Dissertation. Bellaterra: UAB. Retrieved from http://www.tesisenxarxa.net/TDX-1021103-173806/

Moss, H. E., Tyler, L. K., & Taylor, K. I. (2007) Conceptual structure. In G. Gaskell (Ed.), *Oxford Handbook of Psycholinguistics*. Oxford, UK: Oxford University Press.

Müller, G. (2011) *Constraints on Displacement: a Phase-Based Approach*. Amsterdam: John Benjamins.

Patel-Grosz, P. (2012) *(Anti-)Locality at the Syntax-Semantics Interface*. PhD Thesis. MIT.

Putnam, M. & A. Fábregas (2014) On the need for formal features in the narrow syntax. In Kosta, P. et. al. (eds.) *Minimalism and Beyond: Radicalizing the Interfaces*. Amsterdam, John Benjamins. 109-129.

Rizzi, L. (1990) *Relativized Minimality*. Cambridge, Mass.: MIT Press.

(2009) Movement and Concepts of Locality. In Piatelli Palmarini et. al. (eds.) *Of Minds and Language*. Oxford, OUP.155-168.

Saddy, J. D. (2009) Perceiving and Processing Recursion in Formal Grammars. Paper presented at the *Recursion: Structural Complexity in Language and Cognition* Conference at the University of Massachusetts, Amherst, 26–28 May, 2009.

Sag, I. (2007) Remarks on Locality. In Müller, S. (ed.) *Proceedings of the HPSG07 Conference*. Stanford: CSLI Publications.

Schneegans, S. (2014) Operations in Multi-Dimensional Neural Fields. Presented at KogWis 2014, University of Tübingen. http://www.uni-tuebingen.de/index.php?eID=tx_nawsecuredl&u=0&g=0&t=1422113823&hash=488dd323d865d3039c9f2a325a34de6c02e3cdae&file=fileadmin/Uni_Tuebingen/Fakultaeten/InfoKogni/WSI/Zentren/CogSciCenter/tutorialDynamicFieldTheory_Schneegans_NDFields.pdf [Retrieved on 16/01/2015]

Schöner, G. (2009) Development as Change of System Dynamics: Stability, Instability, and Emergence. In Spencer, John et al. (eds.) *Toward a Unified Theory of Development Connectionism and Dynamic System Theory Re-Considered*. Oxford: OUP. Retrieved from Oxford Scholarship Online. DOI:10.1093/acprof:oso/9780195300598.003.0002.

Searls, D. (2002) The language of genes. *Nature* 420. 211-217.

Shirley, E. (2014) *Representing and remembering Lindenmayer-grammars*. PhD Thesis, University of Reading.

Siegelmann, H. (1995) Computation Beyond the Turing Limit. *Science* 268. 545-548.

Spencer, J., S. Perone, and J. Johnson (2009) Dynamic Field Theory and Embodied Cognitive Dynamics. In Spencer, John et al. (eds.) *Toward a Unified Theory of Development Connectionism and*

Dynamic System Theory Re-Considered. Oxford: OUP. Retrieved from Oxford Scholarship Online. DOI:10.1093/acprof:oso/9780195300598.003.0005

Sperber, D. & D. Wilson (1995) *Relevance: Communication and Cognition*. Oxford: Blackwell.

Spivey, M. (2007) *The Continuity of Mind*. Oxford: OUP.

Starke, M. (2011) Towards elegant parameters: Language variation reduces to the size of lexically stored trees. Ms. http://ling.auf.net/lingbuzz/001183

Strawson, P. (1950) On Referring. *Mind*, Vol. 59, No. 235. 320-344.

Stroik, T. (2009) *Locality in Minimalist Syntax*. Cambridge, Mass.: MIT Press.

Talmy, L. (1983) How language structures space. In Herbert L. Pick, Jr. & Linda P. Acredolo (eds.) *Spatial orientation: Theory, research, and application*. New York: Plenum Press. 225-282.

(1991) Path to realization: A typology of event conflation. *Berkeley Linguistics Society* 17: 480-519.

(2000) *Toward a cognitive semantics.* Cambridge, Mass.: MIT Press.

Taylor, K., B. Devereux & L. Tyler (2011) Conceptual structure: Towards an integrated neurocognitive account. *Language and Cognitive Processes*, 26:9. 1368-1401.

Taylor, K. I., Moss, H. E., & Tyler, L. K. (2007). The conceptual structure account: A cognitive model of semantic memory and its neural instantiation. In J. Hart & M. Kraut (Eds.) *Neural Basis of Semantic Memory*. Cambridge, UK: Cambridge University Press. 265-301.

Toyoshima, T. (1997) Derivational CED: A consequence of the bottom-up parallel-process of Merge and Attract. *Proceedings of the 15th West Coast Conference in Formal Linguistics*. Stanford: CSLI. 505-520

Uriagereka, J. (1998) *Rhyme and Reason*. Cambridge, Mass.: MIT Press.

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

(2002a) Multiple Spell-Out. In Uriagereka, ed. *Derivations: Exploring the Dynamics of Syntax*. London, Routledge. 45-65.

(2002b) Warps: Some Thoughts on Categorization. In Uriagereka, J. *Derivations: Exploring the Dynamics of Syntax*. London: Routledge. 288-317.

(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.

(2014) Regarding the Third Factor: Arguments for a CLASH model. In Kosta, Peter, Steven L. Franks, Teodora Radeva-Bork and Lilia Schürcks (eds.), *Minimalism and Beyond: Radicalizing the interfaces*. Amsterdam: John Benjamins. 363–391.

Watumull, J. (2012) A Turing Program for Linguistic Theory. In *Biolinguistics*, 6(2). 222-245.

Watumull, J., M. Hauser, I. Roberts & N. Hornstein (2014) On Recursion. *Frontiers in Psychology: Language Sciences* 4: 1017.

Wilson, D. & D. Sperber (2003) Relevance Theory. In L. Horn and G. Ward (Eds.) *Handbook of Pragmatics*. Oxford: Blackwell. 607-628.

Zeeman, C. (1977) *Catastrophe theory: selected papers 1972- 1977*. Cambridge, Mass.: Addison-Wesley.

Zipf, G. (1949) *Human Behavior and the Principle of Least Effort.* Cambridge, Mass.: Addison-Wesley.