Internal Merge Beyond Explanatory Adequacy

Diego Gabriel Krivochen

Abstract

In this paper I will analyze the conceptual and computational motivations of the property of *displacement* in natural languages from a Radically Minimalist point of view. I will account for *displacement* phenomena proposing my own version of *displacement-as-external token Merge*, as opposed to the traditional *displacement-as-literal movement* or, more recently, *displacement-as-copy and Merge* (Chomsky, 1995; Nunes, 2004). Going *Beyond Explanatory Adequacy* is a notion I will understand here as aiming to grasp the ultimate motivations of phenomena in terms of bare *interface requirements* in the context of a *massively modular mind-brain*.

Keywords: Radical Minimalism, Movement, Syntax-Semantics interface, Free Merge, Type-Token

1. Outlining the framework

To begin with, I will introduce the basic framework within which I will conduct this inquiry. Radical Minimalism is a program which analyses language as a fundamentally physical system. Since language is part of the natural world in a non-trivial way (that is, it not only occurs in the Universe, but is a system that can be assimilated to others in that Universe), as Chomsky and mainstream Generative Grammar have claimed over the years (see Chomsky, 1965, 1986, 2005), it is assumed to be ruled by the same principles as the null hypothesis. The program strives to answer questions concerning the integration of what we refer to as

language in a system of interacting cognitive capacities and furthermore advocates that language as a *physical system* should therefore not be studied in isolation, but rather in interaction with other systems. As this interaction occurs in the so-called 'natural world', it is constrained by physical laws which are in turn particular instantiations of mathematical possibilities. Considering this scenario, Radical Minimalism proposes the following tenets:

(1) Language is part of the "natural world"; therefore, it is fundamentally a physical system.

(2) As a consequence of 1, it shares the basic properties of physical systems and the same principles can be applied (*indeterminacy*, the *Conservation Principle*, *locality*, etc.), the only difference being the properties of the elements that are manipulated in the relevant system.

(3) The operations are taken to be very basic, simple and universal, as well as the constraints upon them, which are determined by the interaction with other systems, not by stipulative intra-theoretical filters.

(4) 2 and 3 can be summarized as follows:

Strong Radically Minimalist thesis (SRMT):

All differences between physical systems are "superficial" and rely only on the characteristics of their basic units [i.e., the elements that are manipulated], *which require minimal adjustments in the* ***formulation*** *of operations and constraints* [that is,

only notational issues]. *At a **principled level**, all physical systems are identical, make use of the same operations and respond to the same principles*[1].

From SRMT, the regularities that have been found regarding the structure of each faculty in the context of a massively modular mind follow straightforwardly (see Katz & Pesetsky's, 2011 "*Identity Thesis*" as an example), as well as the parallels existing between subpersonal systems (i.e., mental faculties) and other biological / physical systems (see Uriagereka, 1998 for some examples and references). It seems relevant to insist on the fact that I am not making a *reduction* of biology (as language is ultimately a biological system, if Chomsky's claim that language is a natural object is accepted) to physics (and that my use of RM in mathematics and physics is ***not a metaphor***), but simply analyzing a *biological* phenomenon in physical terms, as a physical system (in which there is no contradiction whatsoever) and, as such, applying the *tools* that have been devised in physics in the degree that it is possible, and without confusing the *methodological* tools with *substantive* elements. Of course, looking for exact correlates between *any* two fields would be irrational in the *substantive* level (i.e, units of analysis, as Poeppel & Embick, 2005 correctly point out), but I put forth that the *methodological level* has much to tell us, as they are all "parcels" of the same Universe that, I tried to show in our previous work (mainly, Krivochen, 2011b) and will also argue here, are ***identical in a principled level of abstraction***. Therefore, I claim that *computational* properties (i.e., Merge) can and in fact *must* be formulated substrate-neutrally. I will focus on the human mind-brain, but my hypothesis is much less restricted.

[1] Such a pretention of universality regarding physical explanation is somehow reinforced by the advances in M-Theory and, more humbly, on string theory and field theory (Greene, 1999, among others). At this point, however, it is a desideratum, but one that, if pursued seriously, could lead to significative improvements in our understanding of the relations between physical systems.

I claim that there is *only one* generative operation (in the physical world in general, in the mind-brain in particular), namely *Merge*, which is free, "blind" (that is, insensitive to the *characteristics* of the objects it manipulates, I follow and extend the thesis of Boeckx, 2010 that only *format* is relevant) and unbounded, and an operation *Transfer* that provides us with a way of delivering structured information across modules. Merge is an inherently diachronic operation that generates binary-branched hierarchical structures, endocentricity being merely a C-I interface requirement in the case of what has been called the Faculty of Language (FL) (see Chomsky, 2009). Transfer takes place as soon as it can, and this timing is determined by the formation in real time of a fully interpretable configuration in terms of "bare output conditions", what I call a *phase*[2]. Formally, Merge has the form of an unrestricted *concatenation function* that applies to objects in an *n*-dimensional workspace W:

(1) *Concatenation* defines a *chain* of coordinates in *n*-dimensional generative workspaces W of the form $\{(x, y, z\ldots n) \subset W_X \ldots (x, y, z\ldots n) \subset W_Y \ldots (x, y, z\ldots n) \subset W_n\}$ where $W_Y \equiv W_X \equiv W_n$ or $W_Y \neq W_X \neq W_n$.

So far, I have described a computational system whose material instantiation is something irrelevant: wherever there is complexity, even in a trivial sense, there will always be, by conceptual necessity, a generative operation and the possibility of sharing information in the form of structured symbolic representations. This mechanism is "dumb" in as much as it is insensitive to the properties of the elements it manipulates: elements are put together in the

[2] Cf. the orthodox view on Chomsky, 1998, 1999, 2001, 2005a and Gallego, 2010, who define *phases* as the *locus* of feature valuation processes, a view made explicit in Gallego's (2010: 51) "Phase Condition": Uninterpretable features signal phase boundaries. Notice that, in strict terms, the interfaces are completely excluded in this syntacticocentric view.

working area of a determined module because they share a "common format" (Boeckx, 2010), regardless of the characteristics of these elements, and even the characteristics of the resultant object. From this it follows that *there is no "ill-formation" in the syntax, but only at the interface levels*, if, for example, a transferred configuration does not allow the licensing of a given element, or no explicature can be built with the instructions provided by the syntax in the form of a Logic Form (in Relevance Theory's terms, an incomplete propositional form, with unsatisfied referential variables, either sortal or eventive) representation.

In a "free syntax" there are no constraints at all, so there is no point in positing feature-driven operations as they represent a substantive complication of the theory, rather than being the null hypothesis. Merge applies because two whichever objects share a common format and Transfer applies as soon as interface conditions allow it: necessary conditions are also sufficient conditions for Merge to apply in our free system. The role of features, which has been one of extreme importance in Minimalist syntax[3], is questioned, and the very notion of feature is put to test. My argument goes as follows: let us start with a fully interpretable Relational Semantic Structure (see Mateu, 2000a, b for the original presentation; Krivochen 2010b for developments within this framework), built by merging *generic concepts*, a mechanism that has been also used by Boeckx (2010). According to the *Conservation Principle* (Krivochen, 2011b), which we have borrowed from physics, information must be carried along the whole derivational path (i.e., information cannot be erased, but instantiated in a way so that it can be manipulated by the relevant interpretative system), which implies that the concepts will have to be instantiated in such a way that they can be manipulated by

[3] For an exhaustive analysis and references, see Adger (2010), Adger & Svenonius (2009).

the syntax, those concepts take the form of *roots*. So far, no features or procedural instructions are in play, only conceptual primitives. Apparently, features should be added at this point in the derivation, when a semantic object is transformed into a syntactic object. (Un-)Interpretability depends on valuation (Chomsky, 1999), and valuation depends on the category on which F appears. Those features that enter to the derivation unvalued in a category must be eliminated for the derivation to converge. Our objection here is: why adding features in the first place if the system will then eliminate (some of) them? This, without taking into account the stipulation that underlies the whole system regarding the fact that a feature [X] enters the derivation valued in category P but not in category Q. Even if the reader does not accept our use of the Conservation Principle, this second objection is valid within an orthodox Minimalist framework. Feature valuation-deletion also entails the following problem, first noticed by Epstein & Seeley (2002): *the timing of Spell-Out*. If the orthodox view that Spell-Out deletes the features that are uninterpretable by LF is accepted, (i.e., those which have entered the derivation unvalued, and have therefore acted as *probes*, copying the value of a c-commanded *goal*) then the system must be told which of all the features that appear in a certain derivational point had entered the derivation unvalued. But, in order to do so, it would have to look back and see the derivational point immediately before valuation, which is impossible in a derivational (even if it is not as strong as Epstein's 1999) approach as the derivation is a diachronic process, and past states of the system are no longer accessible. The situation can be summarized as follows:

Spell-Out timing:

a) Prior to valuation. Result: *crash*. Uninterpretable features get to the interface levels.

b) <u>After valuation</u>. Result: *crash*. There is no way of knowing which features entered the derivation unvalued (and were, therefore, uninterpretable by LF).

Chomsky (1999) attempted to solve the problem by stipulating that Spell-Out (i.e., *Transfer* to PF) takes place "shortly after" valuation, but I do not see how this could help solving the problem. Epstein & Seeley also tried to provide an explanation by saying that the transference took place *within a transformational cycle*, that is, not before, not after valuation, but *during the process*. For us, that is not a satisfactory answer, since it is simply not principled, but stated. My solution is quite more radical: I just *eliminate features from the picture*. My proposal is the following: instead of binary-valued dimensions, there are only primitive dimensions with procedural value (Case, to give a well-known example) that can adopt a number X of values, in this case, 3: Nominative, Accusative and Dative (Krivochen, 2011b, 2012b). In isolation, the relevant dimension will *comprise all three values* because if all three outcomes are possible states of the system, their linear combination is a possible state if the system as well. The relevant dimension will then be in what I call (following Quantum Mechanics) the *ψ-state*. The ψ-state of the Case dimension would be something like this:

(2) $[\text{Case}_X] \longrightarrow N\varphi + A\theta + D\lambda$

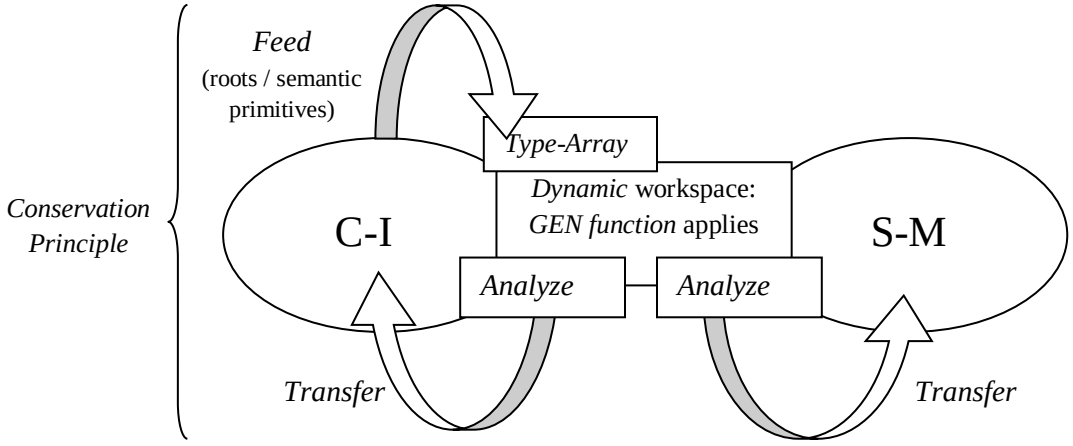
The system can manipulate an element with a dimension in its ψ-state (as it is blind to the characteristics –i.e., "features"- of the elements it manipulates *but* ***format***), but there will be a collapse at LF (again, in RT terms), since the procedural contribution of the dimension will be the same as null as no mental module can interpret ψ-states (see Schrödinger, 1935 for a simple but extremely clear and useful example). That state will hold only if the element is not inserted in a syntactic (i.e., structural) configuration, in which it appears in a Minimal

Configuration (Rizzi, 2004) with a *functional-procedural* node, whose primary function is to determine the relation that is to hold between conceptual elements (i.e., *concepts* or *conceptual addresses*, using Boeckx's 2010 terminology; *generic concepts* or *roots*, in ours). The functional-procedural head in a local relation with the dimension makes it collapse to one of its possible states depending on the "interpretable" procedural semantic features it carries. Therefore, if the procedural head in question is P (which conveys *locative* instructions), to take an example, the closest [$Case_X$] dimension will be *collapsed* to the Dλ state, namely, the *Dative Sphere*. The advantages of this view are clear: there are neither *probes* and *goals*, nor the stipulation that the same dimension be in both. No more *valuation* / *interpretability* problem, no more stipulations as regards where features are or any of the concomitant proposals (feature inheritance / sharing, etc.). Besides, I agree with Boeckx (2010) in that, if the generative algorithm is really *free*, Agree is a(n *unprincipled*) form of constraining it, which goes against Radically Minimalist desiderata. As I already claimed in Krivochen (2011b), the correlations (*Split*)T-Nom, cause-Acc and P-Dat can be derived from independent interface requirements, the dynamics of the derivation in interaction with the interface levels *input* conditions: the conditions an object must fulfill to be read by the relevant system. In our architecture, then, syntax may be blind, but C-I is certainly *not*[4]. The C-I interpretative system, which in massive modular models of the mind takes most of the functions of Fodor's (1983) central processor, not only *drives* syntactic derivations (as the

[4] We will not address in depth here the question whether C-I is computational (i.e., "generative" in some way) or not. Fodor (2000) and Panagiotidis (p.c.) argue against it. We claim (quite in the line of Jackendoff, 1987, 2002 and building on ideas from Mateu, 2000) that there are purely conceptual structures that, even though they *can* be instantiated linguistically, they are not linguistic in nature. Our view clearly contradicts the conception about the semantic component that has been accepted since early Generativism (see Chomsky, 1965, for example), namely, it is *only interpretative*. For us, there is *syntax* in every module that deals with *structured symbolic representations*, therefore, C-I *must* be computational.

Conservation Principle predicts) but it is in itself a *generative* W: complex conceptual structures (of the kind described by Jackendoff, 1990, 2002 for example), are built by *Merge*, in consonance with our claim that Merge is the one and only generative mechanism in the physical world, and thus in the human mind. Of course, it may be claimed that very little is known about C-I (or the Universe in general) to make such assumptions, but what I am doing here is inferring what characteristics it should have if it is the case that *the mind-brain can actually described and explained in terms of Radical Minimalism*. Whether it is or not, is still to be empirically proven. The architecture I argue for looks like follows (taken from Krivochen, 2012b):

(3)



2. External Merge theory of Movement and the *type-token* relation

In this section, I will show that understanding Move in terms of *actual displacement* (as in the GB model) violates the *Inclusiveness Condition* (IC), introducing not only new elements in the derivation, but also a new set of principles to rule dependencies between the moved

element and the trace left behind, which is a different element. In Chomsky (1998) the *Copy Theory of Movement* is explained in length, pointing out that having Copies instead of Traces goes better with theta-theoretic considerations and it would not violate the IC (see also Nunes, 2004 for a clear presentation of the theory). However, there are still extra elements: numerical sub-indexes in the NUM. Let us assume now that there is no NUM (i.e., no numerical sub-indexes), but simply an *Array*, consisting on *types*, of which an (*a priori*) unlimited number of *tokens* could in principle be realized in the syntactic workspace. Let us work with the following set of assumptions:

(4) A *type* is an abstract element in a physical system Φ.

(5) LEX_S is the full set of *type*-symbols that can be manipulated by a computational system S, which is a generative W. An *Array* is a set A of *types* such that $A \subset LEX_S$.

(6) A *token* is an occurrence of a *type* within W_X. There are no *a priori* limits to the times a *type* can be instantiated as a *token* but those required by Interface Conditions IC.

(7) If α and β are interface-associated via their coordinates in W, there exists a *Dependency* between α and β.

(8) If $\alpha \in W_X \wedge \beta \in W_Y$ and either $W_X \equiv W_Y$ or $W_X \not\equiv W_Y$ and α and β are defined by the same *n*-plet of coordinates in their respective *isodimensional* Ws, α and β are *bound* and the *dependency* is called *co-referentiality*.

The property of displacement reduces to *External Merge a token from the type-Array* if we consider that a licensing node can trigger the E-Merge of a token in order to fulfill Interface Conditions, the Conservation Principle (ConsP) and Dynamic Full Interpretation (DFI):

(9) ***Conservation Principle****: Dimensions cannot be eliminated, but they must be instantiated in such a way that they can be read by the relevant level so that the information they convey is preserved.*

(10) ***Dynamic (Full) Interpretation***: *any derivational step is justified only insofar as it increases the information and/or it generates an interpretable object*.

The number of tokens required is determined by interface conditions, so that the *minimal number* of tokens leading to a convergent object is used, provided that the notion of "convergent object" does not arise from look ahead (the syntactic component looking at the legibility conditions of the interface systems, which would lead us to a "bad" crash-proof system, in which a partly interpretative system makes interface requirements superfluous: a *constructivist system* of the kind defended by Lasnik, Uriagereka & Boeckx, 2005 Chapter 2; or Frampton & Gutmann, 2002), but rather from the interfaces "peering into" the syntax (something similar to the proposal by Boeckx, 2007, although the consequences he draws are completely different from mine) and *Analyzing* after every application of Merge whether a syntactic object is ready to be transferred (Krivochen, 2011d, 2012b; Kosta & Krivochen, 2011). A strong version of our proposal allows us to dispense with the very notion of movement in terms of actual "displacement". Let us assume that an *Array* contains *types* of units (whose conceptual or procedural nature will be identified at the semantic interface, this is why I avoid the term *Lexical Array*), not *tokens*, as it is commonly assumed. If this is correct, then I could put forth the idea that there is no upper bound on occurrences of a type in a derivation (i.e., the number of times it is instantiated as a token) beyond what is required by

the *Conservation Principle* and *Dynamic Full Interpretation*. The very elements of the *Array* are instantiations of the elements present in the purely conceptual Relational Semantic Structure (Mateu, 2000a, b) –very reminiscent of Jackendoff's (1987 et. seq.) Tiers-, but they are types, therefore, *no numeric sub-indexes are needed*. There is no movement in terms of actual displacement if one considers that all Merge is external Merge: instead of moving (i.e., displacing) a syntactic object, the system, driven by interface necessity, can simply draw another instance (i.e., token) of that object from the Array and Merge it externally, the process occurring over and over again until no further instances of that SO are required by C-I to build an explicature (i.e., a fully-fledged propositional form with satisfied referential variables, both eventive and nominal, enriching the LF. See Wilson & Sperber, 2003 for details and references). I have thus dispensed with an operation that required stipulative notions for its implementation, and have a simple, free, unbounded generative system instead, which I take to be a desirable result. To summarize, I eliminate sub-indexes from the *Array* and implement a *type-token* relation in a Free-Merge system. By this means, I do not restrict stipulatively or *a priori* how many *tokens* will Merge in structural positions that lead to optimally relevant interpretations in both interfaces. Movement is therefore redefined in terms of *externally merging a token of a type in the Array in a position, licensed by a procedural element, and in which there is an increase in the informational load*, following DFI. The derivational dynamics I argue in favor of is as follows (Krivochen, 2012a: 7)

> "**Definition 15:** *Select* instantiates a *type* in a W_X following **Principle 2** [the Conservation Principle].

Definition 16: *Merge* concatenates LI-tokens in a W_X driven by the interfaces' constraint expressed in **Definition 13**. [a *token* is never fully interpretable at the relevant Interface Level IL unless within a larger structure.]

Definition 17: *Analyze* evaluates the objects built via *Merge* in W_X in order to verify full interpretability in IL_X.

Definition 18: *Transfer* is the operation via which an Interface Level IL_X takes a fully interpretable object from W to proceed with further computations.

Corollary: if W_X interfaces with more than one IL, *Transfer* applies for each IL *separately*.

Definition 19: *Merge, Analyze* and *Transfer* are both interface-driven and interface required.

Definition 20: 15, 16, 17 and 18, occurring cyclically, determine the derivational dynamics in W_X."

Array is not a primitive concept, or anything like a *representation level*: it is never interpreted, and no conditions on well-formedness apply to it: it is simply a finite set of *tokens*. In this sense, my proposal greatly differs from Martin & Uriagereka's (2011) hypothesis involving *types* and *tokens*. They claim:

> "*There is no natural way of capturing tokens, other than by refying lexical arrays into numerations, objects that exist pretty much in the technical sense of a level of representation*"

The reader must have noticed that I have captured the type-token relation in both a formal sense and at the semantic interface, and also accounted for *selection* via Conservation Principle, without resorting to any new assumption, at least within the present framework. Numerations, strictly defined, require numerical subindexes, which are actually *diacritics*, whose interface justification is hard to find. Actually, once one has a natural way to capture

types, *tokens* follow naturally without stipulations in a Free-Merge framework, interface-driven.

The advantages of a *Radically Minimalist* framework are the following, in a nutshell:

- No features, just *configuration* to relate constituents (Cf. Pesetsky & Torrego's 2004, 2007 feature system, or Müller's 2011b rich taxonomy of features)
- Only one operation: (External-token) Merge-α
- Less elements: only roots and procedural nodes (Cf. Exo-Skeletal Models –XSM- like Borer's 2005, 2009 or some enriched Distributed Morphology proposals, like Panagiotidis' 2010)
- No units are syntactically exceptional if *format* is all that matters (e.g., Clitics, as they are analyzed in traditional literature)
- Only *third-factor*, interface-required operations since interfaces are invasive and trigger the application of operations in the syntactic workspace.
- Deeper explanatory adequacy, as we take into account both the biological and computational implications of our claims with the perspective of the interface systems, for a theory of which we offer prospects developing well-grounded ideas with particular interest in cognitive semantics (see Jackendoff, 1983, 2002; Culicover & Jackendoff, 2005; Anderson, 1977; Talmy, 2000; Pylyshyn, 2007, among many others).

2.1 Reconstruction procedures at the Interface Levels and the *type-token* relation

An account of Displacement should address the question of how the interpretative component (i.e., C-I, the inferential module in Relevance Theory) puts together the information that has been provided to it by multiple Transfer? There must be, apparently, some connection between phases, and that connection would be the role of "phase edges" (Boeckx, 2010; Gallego, 2010). However, that only makes sense if one posits that phases are *endocentric* and that labels exist in the syntax. If I do not distinguish between X' and XP, there is no point in talking about edges (i.e., Specs.), at least in the traditional sense. Phase edges are said to have relevance to reconstruction processes, that is, the system can trace back the derivational path of a certain element by looking at its previous positions in the periphery of the phases it had to move through to get to its final destination. In my model, however, things are analyzed differently. To begin with, our phases are not endocentric, but defined in terms of *convergence,* as the ***minimal term*** *in a certain level fully interpretable by the next component* (Krivochen, 2010b). If I combine this with a radically "bare phrase structure" (Krivochen, 2011b), what results is a picture in which there is neither real "phrase structure" (X-bar theory) nor "phase structure" (Boeckx, 2008), but only *structure*. That is, *there are no phrases in the syntax, there is no projection or labeling either* (which is the expected scenario if we take the hypothesis that the syntactic component is just *generative* seriously)*, only free applications of Merge and Transfer of fully interpretable units*. The role of edged and their very existence only make sense if there is something that is *not an edge* (a head or a complement), and that is simply absurd in our framework. Of course there are reconstruction effects at the interface levels, and, in fact, dependencies across phases are *only relevant at the interface levels*, since those dependencies are interpretative, and syntax is a "blind" generative component. Let us put this straight:

(11) If α and β are interface-associated via their coordinates in W, there exists a *Dependency* between α and β.

(12) *Dependencies* are read off in IL, not in W.

(13) A *dependency* is *Local* if and only if there is no intervenient object γ (of arbitrary complexity) such that: (i) the relation between α and γ is equivalent to that between α and β for interface purposes (ii) α, β and γ belong to the same W and (iii) γ is structurally closer to α than β

Therefore, the only requirement for reconstruction to take place is to give the interpretative component some clue that what generativists used to call a *chain* CH = $(X_i \ldots h_i)$ must actually be expressed in terms of dependencies between *tokens* at the relevant interface level. Movement and copy erasing are thus expressible in terms of *multiple occurrences tokens of the same type* of an element in a *type-array*, motivated by "drastic interface effects", that is, effects in the *explicature*. This way a principled explanation for the Spell-Out of *only one* of the tokens, in standard cases, can be proposed: if an element is displaced, the index (in traditional binding terms) should be maintained across the derivational path, and the simplest way to do this is by materializing only the copy whose structural position leads the system to optimal relevance. For example (assuming Grohmann's 2003, 2004 analysis):

(14)

a) John wants ~~John~~ to leave. *Optimally relevant, LF sees both occurrences of [John] as different tokens of the same type, i.e., they are "coindexed" for explicature purposes.*

b) John wants John to leave. *Only optimally relevant if disjunct reference is understood. The Spell-Out of both [John] is interpreted as two different types, and the coindexation cannot take place.*

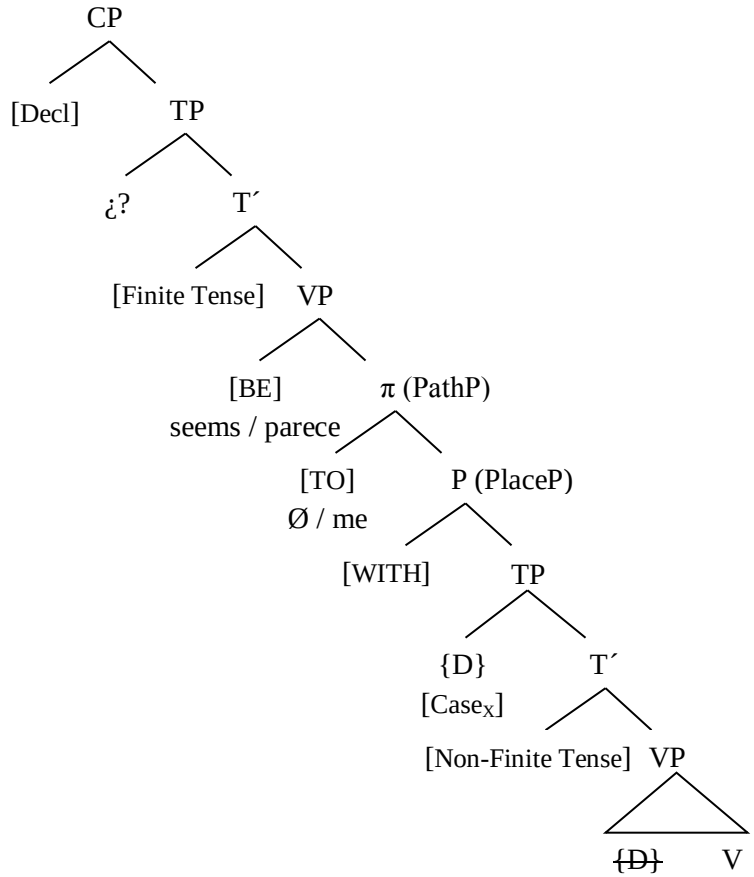
If an element is Spelled-Out, that is apparently more costly than leaving it as a null copy (not in terms of computation, but in more concrete terms of "linguistic machinery"). The generalization would be "if you can leave something covert, do so. If you make it overt, then you must have a powerful reason for that (e.g., you want to generate some positive cognitive effect that could not have been generated with the covert option)", that reason being very much in the spirit of Grohmann's (2003, 2004) *Condition on Domain Exclusivity* "drastic effect on the output". However, even if *both* instances are Spelled-Out, the interpretative system looks for an interpretation, since there is a presumption of optimal relevance that makes the system analyze all the possible interpretations serially until relevance expectations are fulfilled. Copy-erasing would be a clue, like many others (e.g., procedural categories), that leads the inferential system to the intended interpretation. Of course, different languages posit different requirements when it comes to Optimal Relevance at PF, accordingly, the materialization of more than one copy could be required (see, for example, the examples in Nunes, 2004: 38 ff., who attributes the realization of *n* copies to LCA effects). This is not banned by our system; we just state the economy condition that *the minimal number of copies required by PF to achieve Optimal Relevance should be materialized*.

2.2 The analysis put to test

2.2.1 Raising-to-subject movement

In order to provide some evidence for the theoretical claims I have been making, I will address now a problem that has been analyzed from very different perspectives: raising-to-subject movement. I will focus on, but not only refer to, instances of "seem"-type verbs, but, in a more general spirit, the movement of any constituent to Spec-TP. The solution to the problem may have to do with the concept of *theme* (informationally understood, as opposed to *rheme*), and the requirements of the semantic component. Let us consider a typical "seem-raising structure" with a *split spatial domain* (Acedo-Matellán & Mateu, 2010, Krivochen, 2011a):

(15)



What has to be done now is check whether the conditions for Case interpretation of the DP at the semantic interface obtain. Let us proceed bottom-up. After the complete assembling of the thematic domain (*v*P / VP), a T node merges. This T, being non-finite, cannot license a NOM Case interpretation in the relevant argument. If V tried to license ACC / OBJ in that {D} the derivation would crash in the explicature level since (as raising Vs are unaccusative) there is no [cause] primitive that conveys affectedness and can generate a *theme* interpretation, so an appropriate "probe" for NOM licensing must be found. All the nodes that are merged until matrix T are unable to license NOM, so they are not intervenient in terms of Minimality. If the only requirement to license NOM is *absolute*[5] T, which would be the optimal scenario (and the one I will assume), then matrix T is the only appropriate "probe". As there are no intervenient heads between T and {D}, T can generate a Case interpretation for the {D} construction. So how and why is Spec-TP filled? The answer seems to have to do, as I said above, with the concept of *theme*. One proposal could be the following: [EPP] can be dispensed with if the thesis that Spec-TP is an *informationally relevant* position, in terms of the construction of the explicature, is entertained. Of course I am not trying to do discourse analysis here, but it is known that the dynamic *old information-new information* is important for determining the relevance of a proposition[6]. If it only provides old information, it will not be relevant. If it provides only new information, then the Inferential Component (in

[5] Needless to say, it would be false to claim that all non-finite forms convey relative tense, as the morphological expressions do not correspond to semantic interpretation. Arguably, then, Latin historical infinitives convey absolute tense in specific environments, in local relations with other procedural nodes.

[6] For example, Sperber & Wilson (1986a) establish that, *ceteris paribus*, if a proposition provides only old information, it will not be relevant, as no positive cognitive effects (i.e., inferred propositions) will be obtained. If the proposition conveys totally new information, there will be no context to compute it in. One way or another, *optimal relevance* is not achieved.

Relevance Theoretic terms), which I identify with C-I, will not be able to select an appropriate context to process it, and optimal relevance will not be achieved. Elements that move to (i.e., are *remerged in*) Spec-TP are *themes*, and when there are mere presentational sentences, which introduce new referents to the discourse, an *expletive* (either overt or covert) is required:

(16) There is a book on the table
(17) Hay un libro sobre la mesa
Is$_{3SgExistential}$ *a book on the table*
'There is a book on the table'

[a book] is not an element that can be independently manipulated by C-I, that is, it is *not a phase* (at IL) understood as "minimal fully interpretable units", as we will define them here (see Krivochen, 2012b for a development of the argument). Indefinite elements cannot rise because they *cannot be thematic*, as they are newly introduced participants (or are presented as such). Of course, it is not the case that (in-) definiteness is an inherent property of certain morphological realizations of the functional category D, but it is an interface-reading of the cumulative influence nodes like T, Asp and Mod have over an argument (i.e., {D…α…√}, being α an *n* number of non-intervenient nodes for category-recognition purposes, based on specified distribution). Sentences like (16) and (17) are called *thetic sentences*, as they lack *theme*. Spec-TP seems to be a position reserved to elements whose features of definiteness have already been licensed (in the way described in Krivochen, 2010a) and that are thus able to function as *theme*. Violations of these descriptive generalizations rarely generate ungrammaticality (since no principles of the grammar –if there are such things in the first

place- are violated), but semantic anomaly at the *explicature* level, *in the interface*. See, for example, (18):

(18) ?! There seems John to be in the room (irrelevantly, [There seems to be John in the room] is equally anomalous)

[John], as a proper noun, is by definition *thematic*, and thus its raising is "obligatory" for the explicature to be built in order to satisfy relevance expectations. But bear in mind that Case can be interpreted *without raising* (even in feature matching frameworks), and therefore, it would be wrong to regard (14) as "ungrammatical" because of Case reasons, which is what the generativist orthodoxy would say. Movement to the external position licensed by T is, then, triggered by interface conditions, namely, C-I conditions on the construction of the *explicature*.

Bear in mind that the structure above is simplified, since TP is not *Split*, as it should be. Consider a hierarchical structure with separate T, Asp(ect) and Mod(ality) nodes. If token merge in a position licensed by T is motivated by *themehood* (a claim that is too strong and general to be true, but empirical evidence seems to support it at least to some extent if we take, for example, the Spanish paradigm for existential sentences[7]), I can very well assume that movement to merge with each of the nodes is motivated by a different semantic reason.

[7] For instance, if the subject of an Unaccusative V is thematic, it must "raise"-if there is a location-, if it is not (a presentational existential, for example), it can remain in situ. If the subject is thematic but there is no overt location, it can stay, and if it goes up, then this movement targets a Top position. Notice:

a) Llegó Juan. (arrived Juan)

b) ?Juan llegó (if [Juan] is not contrastive or otherwise marked)

c) Juan llegó a su casa (Juan arrived to his house)

d) Llegó Juan a su casa (in River Plate Spanish, only acceptable if [Juan] is contrastive).

Let us take [Mod], for example. Taking into account that the outlined model is strongly componential, there is no way of determining *a priori* what the motivation may be in a given case: the specific derivation in hand must be examined. But, for the sake of expository simplification I can say that, depending on whether the clause under consideration is a simple matrix clause or a subordinate clause, the "subject" of the sentence (in traditional terms) can be the modal subject, that is, the person that makes the modal judgment in either epistemic or deontic terms; or the modal object, that is, the object of the modal judgment by the subject of the main clause. Ultimately, it all reduces to a question of *scope* in LF, as the reader may have imagined. Needless to say, an element "retains" the semantic properties of the intermediate places it has "moved to" (*been token merged in*) in the course of the derivation, so that a {D} structure can be read off at the semantic interface as at the same time a modal subject and the theme of a clause, if there are occurrences of that element in the external positions of Mod and T respectively. We have, then, an elegant account of displacement, without resorting to anything but interface conditions and *explicature* possibilities in the search for *optimal relevance*, which has been already deeply analyzed (Sperber & Wilson, 1986; Wilson & Sperber, 2003; Yus, 2010; Escandell & Leonetti, 2000, among many others).

2.2.2 Wh- interrogatives and the structure of the Left Periphery:

In this section I will try to account for instances of Wh-movement, and revise the structure of the so-called "Left Periphery" (Rizzi, 1997, 2004, among others) under the light of the theory outlined above. Within GB, from Chomsky (1986) on, movement to positions higher than TP is thought to be A'-movement to CP, the projection of the functional category Complementizer. Whereas early Minimalist syntax justified this displacement by means of

p(eripheral)-features, uninterpretable [Wh-] or equivalent that are Checked and Deleted in a Spec-Head relation, so that Movement was an essential part of Feature Checking; later Minimalism, with Matching Theory, allowed features to be valued in a distant (yet local) probe-goal relation and thus Movement was triggered by a different kind of feature: EPP (or, more recently, EF) in C_0. Of course, all this is incompatible with Radical Minimalism as it presupposes substantive complications for the theory, particularly the introduction of new features driving the derivation. I will analyze the characteristics of Wh-movement and try to derive displacement from interface conditions, in order to make it *principled*.

Recall that in Radical Minimalism there are no purely syntactic triggers for operations (i.e., no so-called "formal features" to check), since their application is ruled by a single interface requirement, namely, DFI (which we repeat here for the reader's comfort):

(19) ***Dynamic (Full) Interpretation***: *any derivational step is justified only insofar as it increases the information and/or it generates an interpretable object.*

Given the fact that we have eliminated the feature checking / valuation procedure from the architecture of the cognitive system because of its essentially stipulative character, our only choice is to find an interface motivation for displacement (i.e., *token-Merge*) in different structural positions. To begin with, let us state the following premises that will lead our inquiry on Wh- Movement:

- *Token-Merge occurs only if there is a procedural node that requires so* (restrictive statement)

- *No two targets for token Merge can generate the same interface effect* (anti-redundancy statement)
- *The interface effect is a function of the local relation between a conceptual and a procedural node* (componentiality statement)
- *Extraction filters must be reformulated as interface requirements over dependency establishment. If this is not possible, the relevant filter must be eliminated and their effects, accounted for via independent interface requirements* (anti-stipulation statement)

There are two questions must be asked when considering (Wh-) movement within Radical Minimalism: *Why?* and *Where?*. Whereas previous approaches to the issue have taken into account only *syntax*, I will start from "earlier steps", namely, the *intention* (that is, the "I" part of "C-I") of the speaker and the complex semantic object shaped according to that intention. Bear in mind that I have posited a pre-syntactic instance of C-I, and that is where I will depart from. Wh-elements move because of a *semantic* reason, not because of a *syntactic* reason, given the fact that syntax is free[8]. Therefore, I must find a semantic motivation for Wh-movement, and then analyze what the landing site(s) is / are. Let us take a look at the following examples:

(20) What did you buy?
(21) ¿Qué compraste?

[8] Bear in mind that we are not saying that syntax structures semantics (Cedric Boeckx, p.c.), but that (post-syntactic) semantic structures are a *function* (in the mathematical sense) of syntax, that is, they are read off the structure. Neglecting the possibility of pre-syntactic semantic guidelines is neglecting the "intentional" part of C-I. *Pre-syntactic semantics* ($C\text{-}I_1$) determines the limits of what the syntactic structure can convey (as structure is meaningful in itself) and *Post-syntactic semantics* ($C\text{-}I_2$, RT's inferential module) reads the resulting LF and builds an explicature. Otherwise, there would be no information to instantiate linguistically, *concepts* must come from a non-linguistic pre-syntactic module (as they are needed by other faculties).

What buy$_{3SgPastPerf}$
'What did you buy?'

(22) Who came?

(23) ¿Quién vino?
Who come$_{3SgPastPerf}$
'Who came?'

Those are very simple examples in English and Spanish, but they will do. Both (20) and (21) can be paraphrased as "you bought something. I want to know exactly *what*", and (22) and (23) as "somebody came. I want to know exactly *who*". The first part of each paraphrase is what the sentence *presupposes,* given the following definition of *pragmatic presupposition:*

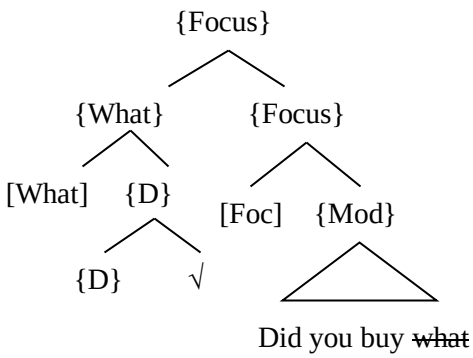
(24) *A presupposes B iff whenever A is true, B is true, and whenever A is false, B is true.* [in other words, B's truth is a *necessary* condition for A's having truth value][9]

That is, Wh-interrogatives generate presuppositions in the semantic interface, as presuppositions are part of the explicature building process. We have claimed in Krivochen (2011a) that Spec-T is a position related to *themehood*, Top generates *entailments* (thus, there is a *contrastive value* in Top-like licensed positions) and Spec-Mod is reserved to the *modal subject*, in modal logics' terms. No presuppositions involved here. We must therefore think of a higher node, semantically interpretable, that licenses an external position for the presupposed constituent, just as [cause] licenses an external position for the initiator. We will call that node *Focus*, maintaining Rizzi's terminology for clarification purposes. Movement to Foc is motivated because the functional-procedural node licenses a remerge of a *token* of the presupposed constituent. We will assume that Wh-elements are Wh- words merged with a {D, √} structure. This is so because D generates an existence presupposition (following Strawson,

[9] This definition is on the line of Strawson (1950) and related work.

1950) of the entity it has scope over, and that presupposition is a *sine qua non* condition for the more global presupposition that is generated (not to mention predication). In other words, if X did not come because X does not exist, the presupposition generated by the Wh-interrogative could hardly exist as well[10]. The structure of the Foc domain in a sentence like (18) would be as follows, using labels for clarity purposes:

(25)



The reader must have noticed that we have not posited T-to-C movement, as Pesetsky & Torrego (2000) –among others- do. The auxiliary [do] conveys [Tense] procedural instructions, thus occupying the "T head" in traditional terms, and the root √BUY receives a "verbal" interpretation at the semantic interface by virtue of being in a [T…α…√] configuration, as we have already seen. The external argument [you] is referentially definite (notice the perfective aspect in the V, which favors EQ interpretations and, in this case, also puts an upper limit to the implicature), therefore, a token is merged in the external position of T to be interpreted as *theme*. However, it need not be Spelled-Out there if the Spell-Out of a structurally lower token leads to plain convergence. This implies that the "*always Spell-Out*

[10] This problem has been analyzed by Russell, on the one hand, and Austin, Quine and Strawson, on the other, with the famous example of the bald King of France.

the higher copy" desideratum that underlies the Copy Theory of Movement (sometimes overtly, sometimes covertly) should be abandoned in favor of a freer system that, moreover, could increase descriptive adequacy (i.e., account for inter-linguistic differences regarding word order) without adding stipulations.

The same derivation can be posited even for *v*-to-T languages like Spanish, the Aux is just replaced by a full heavy verb. As no EPP is involved, no position *needs* to be filled stipulatively. The (labeled) structure would look like:

(26) [$_{Foc}$ [$_{Qué}$ Qué] [Foc] [$_{Mod}$ [$_{Mod}$ compraste] [$_{Asp}$ [$_{Asp}$ ~~compraste~~] [$_{T}$ [$_{T}$ ~~compraste~~] [$_{v}$ [$_{v}$ ~~compraste~~] [$_{V}$ [$_{V}$ ~~comprar~~] [~~$_{Qué}$ [Qué] [$_{D}$ Ø~~]~~]]]]]]]]

There is no external argument, as positing a *pro* null subject there would add nothing to the processing in informational terms, that is, it would not be legitimated by the interface since all the person-number information the corresponding interpretative interface can retrieve from the verbal ending (see Krivochen & Kosta, in progress). The sentence is thus *thetic*.

Notice that I have put a *token* of the V in each of the nodes of Split TP so that it "absorbs" the procedural instructions of the nodes it is e(xternally)-merged on, but actually there is no need to "move" it, I can just posit (following Halle & Marantz, 1993 and subsequent work) that *morphological fusion* has occurred and T, Asp and Mod nodes have been Spelled Out by a single Vocabulary Item, [-ste]. Of course, the selection of the copy to be Spelled-Out is not principled, but depends on *Optimal Relevance* in a global level and *language-specific Spell-Out patterns*. This way, some features and parameters that have proven problematic and *ad hoc* can be dispensed with.

I have provided an account for a traditionally called "discourse-driven" process like elicitation of information from a syntax-semantics interface point of view. Regarding intermediate landing sites, the question whether there is any could arise. As a matter of fact, *token-Merge* applies when [Foc] enters the derivation, crucially not before since there is no interface-driven trigger for the operation. Lower phases could perfectly be transferred as there is no uninterpretable symbol (here, *v* collapses [$Case_X$] in the internal argument to the ACC sphere), and Spell-Out of tokens is driven by *optimal relevance*. In this case, given the Spell-Out patterns of both English and Spanish (which escape the scope of syntactic study, as the B-List and phonological patterns in general depend on sociohistorical matters rather than sub-personal factors) the most relevant token to be materialized is the highest one. In principle, there is no constraint to the number of copies that can be Spelled-Out, but the explicature will either change or require too much effort to process, which can be seen in (27a) and (27b) respectively:

(27)

a. John hit Peter and John kicked Peter
b. What did you buy what?

The only possibility for (a) is interpret disjunct reference for all proper NPs, otherwise, Optimal Relevance will not be achieved: entities are multiplied way beyond necessity, and thus cognitive cost overwhelms positive cognitive effects (i.e., inferences). Apparently, a PF generalization that can be made, building on Chomsky (1995) is that *the minimal number of copies of the smallest syntactic objects leading to convergence must be Spelled Out*, the position being determined by themehood, presupposition, and other semantic reasons. (b) is

straightforwardly too costly to process, as there is no positive cognitive effect that justifies Spelling-Out both tokens of [what]. Movement and Spell-Out have thus been justified in *interface terms*. We will return to this below.

2.2.3 Adjuncts and extraction opacity:

In this section we will deal with a topic that has been analyzed from many perspectives: extraction of material from so-called "adjuncts". In GB, such an operation was banned because of the *Condition on Extraction Domains*, which only allowed extraction of material from complements. In more recent terms, Uriagereka's (1999) distinction between *monotonic* and *non-monotonic Merge* can be summoned to describe the observed effects: if adjuncts, as well as complex specifiers, are generated via non-monotonic Merge, then they belong to a parallel derivation, and therefore are opaque for extraction, but not for movement as a whole. This is the line we will pursue, refining Uriagereka's theory of parallel derivational "cascades" with our mathematical formalization of syntax and, moreover, unifying the semantics of adjuncts so that they can be defined in syntactico-semantic terms without resorting to stipulations or extra elements like A-structure or θ-grid, whose encoding terms have always been problematic (and even more so in a biologically-oriented framework).

Let us start by defining a workspace in Radical Minimalism:

(28) W is an *n*-dimensional *generative* workspace. Taking two distinct workspaces W_X and W_Y, either

a) $W_X \equiv W_Y$ *iff* $\forall(x) \mid x \in W_X, x \in W_Y \wedge \nexists(x), x \in W_X, x \notin W_Y$

b) $W_X \neq W_Y$ *iff* $\nexists (x) \mid x \in W_X, x \in W_Y$

c) $W_X \cong W_Y$ *iff* $\exists(x) \mid x \in W_X, x \in W_Y$

Once we have made explicit what we understand by W, we will work with the following hypothesis, to be refined and enriched later:

(29) *No extraction of α targeting β is possible from inside an element γ (γ ∋ α) if β and γ are assembled in separate W.*

That is: each derivation in each W has its own *type-array,* such that there is one array *per* W. Moreover, both the displaced constituent and the target position (the licensing position) must establish a local dependency. This is essential for our argumentation, since workspaces are instantiated as atomic elements for the purposes of further computations: a clear example is the derivation we have proposed for *path of motion* constructions incorporating a complex root, even though that root is the result of syntactic computations taking place in a separate W. Now, if an element generated in W_X is atomic for the purposes of computations at W_Y, which seems to be a plausible theory, all that is left is justify why so-called "adjuncts" are generated in a different W, and can therefore only be treated as syntactic atoms. The answer, for RM, lies in the *Localist Theory*.

2.2.4 On the localist nature of adjuncts:

Let us analyze the semantics of adjuncts from a localist point of view, taking three main types of adjunct clauses:

(30)

a) Temporal: [Before going to the match], John had a beer.

b) Causal-consecutive: [Because John saw Mary with another man], he left her.

c) Locative: John works [in New York]

Our proposal will be the following: all adjuncts are locative in nature, as there is always an anchoring of an event with respect to a ground, ultimately expressible in Spatial terms.

Let us develop the idea. Assume we have a set E of events (linguistically represented as *v*/VPs), such that E = $\{e_1, e_2, \ldots e_n\}$ and a set T of points in time such that T = $\{t_1, t_2, \ldots t_n\}$. Now, E is an extensional set of *figures*, and T, an extensional set of *grounds*. Should we accept this, the definition of the semantic contribution of different kinds of "adjuncts" can be made explicit in the following way:

(31) A relation $\mathbb{R}$ that holds for the pair $((e_x, t_x); (e_y, t_y))$ is *causal* iff the fact that (e_x, t_x) obtains is a *sine qua non* condition for (e_y, t_y) to obtain and t_x precedes t_y.

This is nothing new, as far as the formalization is concerned. The novelty is to see the relation as *locative* and, moreover, to order the semantic effects of adjuncts in a *scale*. Let us take the *causal* relation as the most complex, since it requires:

(32)

a) Statement of the placement of an event in the system of coordinates taking into account figure and ground (i.e., both sets).

b) The requirement of e_1 to obtain for e_2 to occur.

The leading idea is to simplify the system by recognizing the semantics of each type of adjunct. So, for example, *temporal adjuncts* require describing e_x as a full set of coordinates in an *n*-dimensional space, and then to relate those coordinates to the set defining t_x, thus creating a dependency between them. If this characterization is on the right track, then we do not need (b) to define a temporal adjunct, thus getting a subset relation. Moreover, locative adjuncts only require the explicitation of the coordinates of e_x in W, but no further data. The relation can be represented, then, as follows:

(33) *Causal > Temporal > Locative*

Locative adjuncts require a further note: location can be *literal*, as in the example, or *metaphoric*. In any case, *locative* relations between events can be expressed as "e_1 occurs within the sphere of e_2", as in our Case Sphere Theory. Consider examples (34-35) below:

(34) John was talking *about* his girlfriend.

(35) Juan estaba hablando *sobre* su novia

John was-talking about his girlfriend

'John was talking about his girlfriend'

In these terms, the complement of the preposition indicates in the sphere of what the event takes place. Not unrevealing is the fact that the preposition that is normally used in English, [about] is also used to indicate approximate time, a long-known metaphor of space in cognitive terms (Talmy, 2000; for example). In Spanish, the preposition [sobre] is eminently spatial, as examples of the kind of (36) demonstrate:

(36) Juan puso el libro *sobre* la mesa

John put$_{3SgPastPerf}$ the book on the table

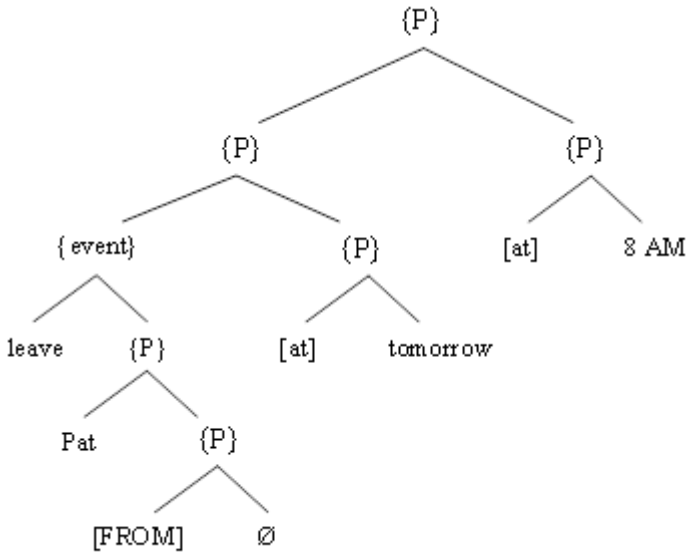
'John put the book on the table'

Of course, inference relates those types, and the characterization is not so clear-cut in real-life examples. However, having a formal theory of the semantics of so-called "adjuncts" is the basis for understanding the inferences, since they are always licensed by elements present in the syntactic structure.

If the semantics of adjuncts is eminently locative, then it is to be expected that it can be analyzed from a *figure-ground* formal framework. I will offer evidence in favor of this approach. In my vision, an event modified by an adjunct is a figure positioned in some relation to a ground, so that (37) is to be reanalyzed as (38):

(37) Pat leaves tomorrow at 8 AM

(38)



In (38), the *figure* in the locative structure is an event itself, and the *grounds* (nothing limits us to only one, neither neurocognitive evidence and let alone the data) the locations in different coordinates of the space-time continuum, conceptualized as an *n*-D space. In this logical space (reminiscent of Wittgenstein's *logische Raum* –paragraph 1.11-, but including a biological dimension absent in the *Tractatus*), multiple workspaces activate, to simultaneously derive, by monotonic Merge, each of the cascades that compose the complex (non-monotonic) structure in (38), namely:

(39) a. {leave {Pat {P, Ø}}}

b. {P, Tomorrow}

c. {at 8AM}

If, as I have said before, Uriagereka's (1999) take on derivational cascades is combined with my hypothesis that extraction out of K can only target Σ if K and Σ belong to the same W at the derivational point T in which extraction is triggered by the introduction of a relevant node in the workspace, so-called "adjunct opacity" follows naturally, and with the added advantage of counting with the support of Talmy's (2000) and Phylyshyn's (2007) localist theory of the mind.

3. Locality and successive cyclicity revisited:

The account of interface-driven displacement I have presented has some similarities with Grohmann's (2003, 2004) *Prolific Domains* approach. In our system, the notions of *locality* and *anti-locality*, of frequent appearance in recent works on syntactic theory, follow naturally from the way in which the dynamics of the (syntactic) derivation interact with interface conditions. What I would like to do is *justify* these concepts, go further than descriptions and

explanations that have been given so far (Rizzi, 2004, Grohmann, 2003, Boeckx, 2010, to name just a few of the most important works). The relevant element would not need to stop in each available position, but only on those that, once the relevant head is merged, the informational load for the interfaces increases, as the derivation proceeds *in real time* and no object can "see" what will be merged later on. This is **why** there is ***locality***: although it can be argued that it would be simpler to wait until the whole derivation is completed and just move the element to its final position, that option would imply maintaining a larger structure in the working memory, a scenario that is far from optimal from a biological point of view (and would wipe phases clean, incidentally). *Successive cyclic movement* is a way of minimizing computational cost. Movement applies as soon as it can (i.e., as soon as a new informational domain is created and there are therefore new heads that can enter in a *licensing* relation with the object in question), waiting would imply departing from the best option, which has to be justified independently. ***Anti-locality*** derives from the fact that each interface-determined domain provides different information to the interface to build a representation, so *intra-domain* movement would be trivial, with no effect in the interface. Informational domains are both *homogeneous* and *heterogeneous* in nature in the following sense: they are *homogeneous* because they are defined taking into account the ***type*** of information they provide the interface with (e.g., both T, and Asp give definition, reference to the generic event denoted by the *v*P domain), and this homogeneity can be seen in instances of *morphological fusion*, heads belonging to the same domain can be fused, as it happens in Spanish with T, Asp and Mod, spelled-out as a single affix (i.e., by means of a single VI). Their *heterogeneity* comes into play when considering these domains closer: all the relevant nodes within a domain convey the same ***type*** of information, but not the same information by any means: if it were the case,

we would have a redundant system. Consider, for example, that V and *v* differ in the specific information they convey: eventivity and causativity, respectively.

It has to be taken into account that **optimally**, *at every point in the diachronic derivation*, interpretability should be satisfied (of course, a rigid interpretation of this claim would lead to a crash-proof system like Frampton & Guttman, 2002, which I do not want. See Putnam, 2010 for an alternative view on crash-proof syntax). That is, if a determined point in the derivation is described as a structural configuration Σ in which a syntactic object SO does not receive an interface interpretation or the informational load could be increased by creating a new object Σ' in which SO establishes a relation with a certain procedural node, this SO *must* "move" immediately, which we will understand in the sense of *monotonically merging a token of the minimal structure containing the relevant element* (optimally, *only* the relevant element), in consonance with the *extension condition*. In a more general spirit, let me say that there is only one "constraint" (in an OT sense) to the application of syntactic operations, *Dynamic (Full) Interpretation*[11]. The reader must take into account that this does not mean that I am changing a strong derivational model for a representational one, since DFI must be satisfied *in real time*, after every application of Merge, as I have said. I strongly stress the *diachronic nature* of Merge. In a representational model (GB, for example); conditions on good formation were applied to fully-fledged representations, namely, D-Structure, S-Structure, LF and PF. My proposal regarding syntax is that it is *free*, conditions upon it are interface-driven and there is nothing more to it. A note of some importance should be made

[11] I will make use of DFI to analyze the topic of *edges*, but it is very useful to gain insight into other areas as well. For example, the analysis of Merge in Adger (2011) can be revisited under this light, as well as the derivation of roots in De Belder & Van Craenenbroeck (2011). See Krivochen (2012c: Chapter 1) for details.

here: I have already argued against *Phase Edges* in previous works (Krivochen, 2010b, 2011b), and here I will provide some more evidence in favor of this proposal. If the claim that an operation is only legitimate if it follows from DFI (that is, interface requirements) in accepted, then a question arises: what (if any) is the difference in legibility between (a) and (b)?

(40)

a) $[_{v\text{P}}$ YP $[_{v'}$ $[v]$ $[_{\text{VP}}$ $[_{\text{V'}}$ [V] $\text{XP}_{[\text{u-F}]}]]]]$

b) $[_{v\text{P}}$ $\text{XP}_{[\text{u-F}]}$ $[_{v'}$ YP $[_{v'}$ $[v]$ $[_{\text{VP}}$ $[_{\text{V'}}$ [V] ~~$\text{XP}_{[\text{u-F}]}$~~$]]]]]$

Let us assume that the "uninterpretable feature" in XP cannot be valued by *v*, but by some other higher head, say, T, or C. Does movement to outer-Spec-*v*P value and erase the feature in question? Certainly not. Therefore, there is an operation that applies to Σ, a non-fully interpretable unit, and generates Σ', an object which is equally uninterpretable by the interface levels. A superfluous operation, put differently. Not only does this take us away from strict *last resort* conception of movement (since the derivation is by no means "saved"), but it also complicates the system both descriptively and explanatorily. Such a theoretical apparatus needs stipulations that I will try to eliminate. My objection, however, is only valid if one accepts *Dynamic Full Interpretation* Principle, which I think of as an important economy principle in a strong derivational approach: I keep syntax blind and free, and attempt to provide a justification for legitimate operations in interface-legibility terms only. Let us see what a derivation would look like in Radically Minimalist terms, starting with the SO

depicted in (a) (I will use labels just for expository purposes, but beg the reader to think of these representations as bare structures):

(41) $[_{v\text{P}}$ YP $[_{v'}$ [*v*] $[_{\text{VP}}$ $[_{\text{V}'}$ [V] $\text{XP}_{[\text{Dx}]}]]]]$

Once this stage in the derivation is reached, there is an element that, if transferred, would cause the derivation to crash as it bears a dimension [D] in its ψ-state. As I have said before, local relation with *v* cannot generate an unambiguous interface reading, suppose it is a feature related with definiteness. In Krivochen (2010a) I -hope to have- demonstrated that definiteness in DPs depended on a local relation with three nodes, Asp, T and Mod, each of which provided the inferential module with different clues for interpretation. So, I will assume this scenario as an example. The next step would be the merger of another element, following *monotonic Merge*. The *v*P domain denotes a *generic caused event*, so the next relevant head would have to be one that can contribute to the specification of this reference. In other words, a set of procedural dimensions extending the structure and increasing the informational load (dimensions that could, but do not need to, be later Spelled-Out) merges following DFI. Let us call this set a T node, for simplification purposes. As the structure is not interpreted by the interface system as a generic event anymore, but as an event situated in a specific time frame (of the form $<e_x, t_y>$ in a traditional Cartesian representation), T has scope over *v*P, and a "label" is recognized by C-I. Remember that labels do not exist in the syntactic Generator since they are irrelevant to it, they are "recognized" (*not* created, so as not to violate the *inclusiveness condition* in traditional terms, or just because they are not necessary in the syntax, in a more contemporary spirit) in the interface level. The same happens with [Asp], a node that conveys the way in which the speaker sees the previously

defined event: as a point or as a developing event, with internal structure. The {Asp} "projection" (AspP, in traditional terms) closes the proposition, since the next node, [Mod], conveys a subjective evaluation over the propositional content (in epistemic or deontic terms), in order to do which this node must not only be outside the proposition but have scope over it. In this structure (in which I have omitted the spatial projection in order not to complicate things) there are two *informational domains* (for the time being), namely, [Mod [Asp [T]]] and [cause [event]]. Informational domains can be indentified because their nodes can be *fused* in Vocabulary Insertion, as I have already said.

The present proposal has direct consequences for the theory of *escape hatches*: in our model, such a notion is simply *unformulable*. Escape hatches (outer-Specs in phase heads) are superfluous positions, which, as I have pointed out, are not justified in interface terms. Besides, the bare structure I have posited in Krivochen (2011b) renders edges unnecessary, as the notion of a Spec position out of a phase domain is only formulable when there are, in the first place, *endocentric phases*, against which I have already argued; and in the second place, projections in the syntax (i.e., X-bar trees, and the concomitant LCA). If there are no phase heads, a direct consequence of defining phases exclusively in terms of minimal convergent units, the notions of *domain* and *edge* lose their theoretical weight. Besides, the mere term *specifier* makes sense when there is a projection system in the narrow syntax, which is actually no more than a labeling algorithm (XP, X', X_0). If syntax can do without *labels*, that is, if labels are *interface-identified* in the explicature level, then the whole phase system has to be revisited critically. Another consequence of our proposals is that the very notion of *successive cyclicity* must be revisited, as the "intermediate landing sites" must be legitimized

by interface requirements, that is, there cannot be an element remerged in a position in which it is superfluous (for details, see Krivochen & Kosta, in progress). In this sense, my proposal is very much in the line of Grohmann (2003), but I tend to a much simpler form of Minimalism. Architecturally, the resultant model is surprisingly simple, but, as the reader must already have thought, there is still a long way to go before full operationalization.

4. Informational Domains and Prolific Domains:

It is time to devote a section to the analysis of what I have called "***Informational Domains***" (ID) in Krivochen (2011b), and compare them with K. Grohmann's (2003, 2004) ***Prolific Domains*** (PD), to see in which ways they are similar and in which ways the operationalization of the ideas differ substantially and the influence these conceptions have for a more general theory of displacement.

First, let us define what Grohmann means by "Prolific Domains" (Grohmann, 2004: 212):

"*A Prolific Domain is a contextually defined part of the computational system, which provides the interfaces with the information relevant to the context and which consists of internal structure, interacting with derivational options".*

CP, TP and VP are thus "prolific" in the sense that they contain more than one "layer":

VP is expanded in *v*P and VP since Larson (1988) (see also Hale & Keyser, 1993).

TP is expanded in Agr_SP-TP-Agr_OP in Pollock (1989) and Chomsky (1995).

CP is expanded in ForceP, TopP, FocP and FinP since Rizzi (1997).

Within *v*P and its associated projections *thematic relations* are established, so this PD will be called *Θ-domain*.

Within TP and its associated projections, *agreement features* are checked (an operation that relies very much on phi-features), so this PD will be called *Φ-domain*.

Within CP and its associated projections, lastly, *discourse-driven* operations are performed. This domain will be called *Ω-domain*.

The revised structure, an expanded version of the classic Chomsky (1986) clause structure is as follows:

(42) $[_{\Omega\Delta}$ ForceP…TopP…FocP…FinP $[_{\Phi\Delta}$ Agr_SP…TP/IP…Agr_OP… $[_{\Theta\Delta}$ *v*P…VP]]]

Grohmann defines them as providing interfaces with information, each domain of a different kind. The informational value of the Ω and Θ domains is clear, and I cannot add much to what Grohmann has said about it. However, the contribution of the Φ domain either to the explicature or to phonology is another matter. Within the Φ-domain, agreement properties are licensed, but I have said that a system with Agree is a stipulatively constrained system, that departs from our idea of a radically simple syntax. Even replacing the notion of "agreement" with pure "licensing" under the scope of a functional-procedural node, the composition of this domain is conflictive from an interface point of view: are AgrS and AgrO fully interpretable? Why would they be needed, in the first place? I have suggested (see Krivochen 2012c, for example) that Case can be accounted for with a system containing only *v*, P and T, any other nodes are superfluous. And, as such, they must be eliminated.

Apart from these technical issues, which can be solved easily even within a PD framework, there are some other aspects to consider. I will analyze particularly two: the *type of information* handed to the interfaces and the *transfer point* ("timing", so to speak).

a) **Type of information**:

Whereas PD handle information about *thematic relations*, *agreement features* and *discourse-driven processes*, I believe that the *theoretical status* of the Φ-domain must be analyzed deeper. Actually, there are no restrictions whatsoever regarding conditions nodes must fulfill to form each PD; as a consequence, Grohmann's clauses include Agr_S and Agr_O, both of which were eliminated in Chomsky (1995) for being superfluous projections for FIP purposes. We posit that if a domain is to be interpreted by the interface systems, then all the projections contained within that domain must be fully interpretable (thinking in traditional X-bar terms). PD seem to be defined in quite aprioristic terms, in spite of the definition Grohmann gives, which appears to be dynamic…and it really is, if it were not for the fact that projections are arranged in advance. The diachronical aspect of the syntactic derivation is therefore lost, and the definition loses much of its explanatory potential.

There is yet another problem, and it is the *interpretability of the Φ-domain*. It is not clear that "agreement" is interface relevant, especially if interpretative systems can rely on structural configuration for establishing relations between procedural heads and root-based structures (i.e., lexical items). It is essential to bear in mind that in no point does Grohmann speak about Tense / Time features, but the Φ-domain is the locus of Agr. As I have done away with *Agree* as it is conceived of in traditional Minimalism, a Φ-domain formulated in Grohmann's terms is neither necessary in or compatible with our proposal.

My *informational domains* operate in quite a different way. There is a single combinatory operation that concatenates elements (say, roots and functional-procedural nodes) following interface requirements. I will also work with three domains, not structurally but purely interface-defined. In other words, there is no place for *a priori* boundaries in our system, consequently, *concatenation* builds a structure and then the interface levels *recognize* domain boundaries *on-line* as fully interpretable material is transferred. Interestingly enough, domains in LF need not coincide with domains in PF: domains in LF are defined taking into account the information they provide for an explicature to be built. In PF terms, if certain nodes can be *fused* into a single VI, they are recognized as a domain (see Krivochen, 2011b):

- {D, {P, D}}: LF: locative information

  PF: no fusion is possible, as the P domain is the one in the bottom, there is no node to fuse with when the P head enters the derivation.

- {cause, {event, {P}}: LF: event including a spatial relation

  PF: V and *v* are fused[12]. P can conflate onto V if and only if there is no lexical insertion, but I would not be talking about *fusion* but *conflation.*

- {Mod, {Asp, {T, {cause}}}}: LF: in strict dominance order, there is a modalized proposition, which includes an event delimited in time and seen either as a point or as developing.

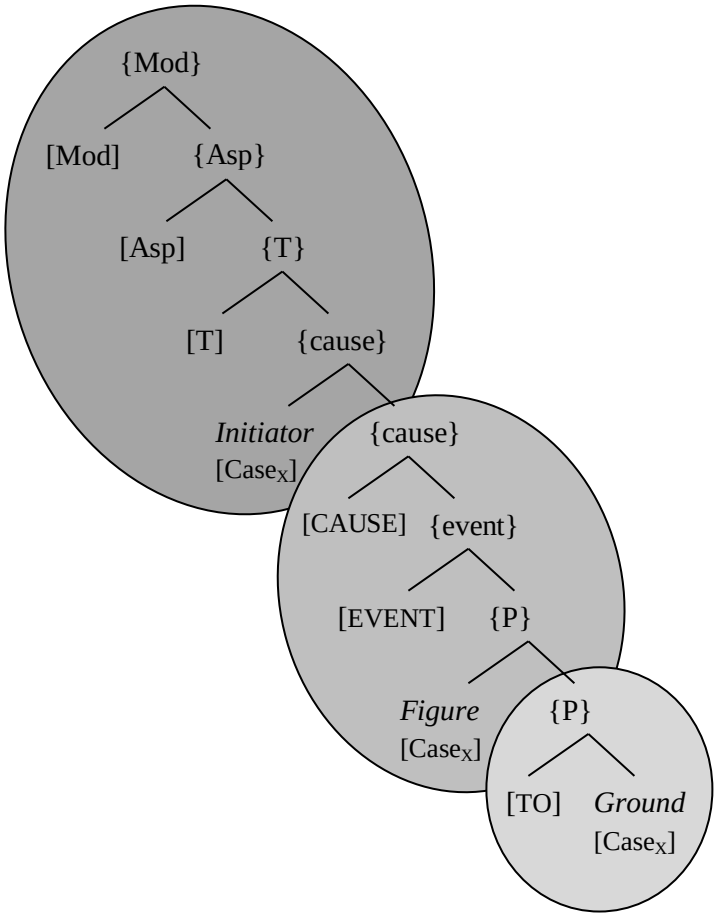
[12] This shows clearly in languages where agentivity is presented as a *morpheme* in the V.

PF: especially noticeable in synthetic languages like Spanish, in which Mod, Asp and T information is Spelled-Out by a single vocabulary item.

A note has to be made here: the fact that in a given language (English, for example) fusion is not possible is a mere historical accident, product of the lack of a single vocabulary item specified enough to spell-out those nodes. This does not mean that the fusion is not *conceptually possible*. Our claim is that the fact that a group of nodes convey the same *type* of information makes the fused spell out possible. We are consciously giving pre-eminence to LF (C-I) over PF (A-P), as there could not be human language without a conceptual interface, but phonology is not only accidental (i.e., not essential, neither necessary nor sufficient) but also *dispensable*, if *communication is an* ***epiphenomenon***. In spite of what some linguists try to put forth, there is no conceptual problem (and less so in a Radically Minimalist framework) in imagining a natural language *without phonology* at all: Merge would only be constrained by C-I requirements. It is very important to bear in mind that, if the *great leap forward* is the emergence of *conceptual addresses* (or *roots*, in our terms), C-I is not dispensable if language is thought of as related in a more general spirit with the infinite use of finite media -echoing Humboldt's words- since C-I *provides the media*, as opposed to the conception of language as a communication-oriented external phenomenon (e-language, in Chomsky's terms).

To summarize, let us remind the fully fledged clause skeleton I proposed in previous works (Krivochen, 2011a, 2012b):

(43)



Labels are included just for the sake of clarity, but I refer the reader to the above discussion and Krivochen (2010b, 2011b) regarding the status of labels. Instead of a labeled tree, the most neurologically accurate representation would be a bare *n*-dimensional (according to recent neurological research, 3-D, but higher dimensions are conceivable even though nonexistent in the phenomenological world: with some effort, anyone can think of a *hypercube*) hierarchical structure, since labels (if interpreted as instructions to take a phrase marker as a unit for further computation) are of no use to syntax: syntax is a free unbounded purely generative mechanism that combines elements with the same format (Boeckx, 2010, Krivochen, 2011b). A Radically Minimalist theory should do its best to dispense with the *copy* operation that is required for labels to appear (and have relevance) in the syntax. Let us

suppose that the following object has been created via Merge: {H, XP}, and is now to be labeled. Chomsky's labeling "algorithm" (it is rather a rule) predicts that H will project. For that to occur, H's categorial features must be copied upwards (an operation that was called "percolation" in earlier models), so that the information is maintained or rather increased: otherwise, there would be no point in labeling at all. Of course, that system presupposes that elements enter the derivation already categorized, a position I have argued against before. If there is no procedural category to collapse the categorial quantum dimension in H, there is no possible labeling, and there is nothing dominating H in the configuration I have analyzed. We put forth that "labels" are *recognized* rather that *created* in the LF interface level, as they are of essential importance for the construction of an explicature. The *inferential module* recognizes informational domains, delimiting them in the most accessible LF for the construction of an explicature.

b) **Transfer point**:

Even though Grohmann explicitly states that a PD framework does not necessarily entail Multiple Spell-Out, I will assume the MSO variant of the PD theory, as it will be useful to compare it with our own version of *derivation by phase*.

Let us first compare the Spell-Out timing of Chomsky's *phase theory* according to the two versions of the Phase Impenetrability Condition (1998; 1999, 2001, 2005):

- PIC_1 (MI): In the phase α, with a head H, H's domain is not accessible to operations *outside α*, only H and its Edge.

- PIC_2 (DbP, BEA, OP): In the phase α, with a head H, The domain of H is not accessible to operations ***at ZP*** [the next strong phase head], but only H and its edge.

Let us assume the following derivation, where H and Z are strong phase heads (say, *v* and C):

(44) [ZP… [XP… [HP [H YP]]]]

Under PIC_1, the transfer of YP occurs when H is merged, as it is explicitly said that no probe outside HP can have access to it, not even XP, which is not a strong phase head (in the terms of Chomsky, 1999). Under PIC_2, however, XP can have access to YP, which means that it is not transferred until Z is merged. Of course, if phases are considered to be the locus of feature checking and transfer, H's condition of strong phase head renders superfluous, as it does not trigger transfer. The "reduction of the computational burden" argument is significantly weakened, as the derivation is transferred when it is completed. The scenario is complicated even more if the relevant object is, for example, a CP appearing as a complement of a transitive V. By considering PIC_2, it is conceptually possible that V can have access to C and its whole domain. The only way to prevent this from happening is to stipulate, for example, that CPs always trigger transfer, which is of course untenable within Radical Minimalism.

We will turn now to PD and transfer point. According to Grohmann (2004), "Spell-Out applies *at* each Prolific Domain –as soon as a PD is formed it spells out." There are no edges here, therefore, if the following configuration is assumed:

(45) $[_{vP}$ ZP *v* $[_{VP}$ V XP]

There can be seen a complete Θ-domain, within which thematic relations are "created" (using Grohmann's terminology, with which I do not agree since relations are *recognized* at the interface, not *created*: this would be a violation of inclusiveness constraints). As soon as *v*P is closed, it spells out *entirely*. This means that there are no edges left behind as in a Chomskyan model. However "natural" this Spell-Out system may seem, it is not stipulation-free. Even though PD provide the interfaces different types of information, and it would seem optimal that each PD be transferred as soon as it is complete, I will try to show that a dynamic definition of a Multiple Spell-Out (MSO) system is "more minimalist". The bigger the structure, the more it is divided (Boeckx, 2010). Having a static system means that there is no relation between the complexity of the structure and its chunking, an option I consider far from optimal, as it does not reflect (describe) or explain the mental mechanisms put in practice. It assumes an automatic mind-brain, which chunks information according to some pre-existing stipulative algorithm.

RM's version of MSO is a dynamic, system-neutral and non-stipulative one. Let us give the definition and then analyze the consequences of adopting such a framework:

(46) *P is a* ***phase in a system*** $\boldsymbol{S_X}$ *iff it is the* ***minimal term fully interpretable in*** $\boldsymbol{S_{X+1}}$

I will maintain Chomsky's (2005, 2007) claim that the *phase* is the locus of *transfer*. Not of feature checking / valuation, of course, since I have dispensed both with features as they were conceived in traditional Minimalism and checking operations. Besides, as our definition applies to any symbolic representation in the mind-brain, it would be unnecessarily restrictive to mention specifically linguistic elements in the definition: *transfer* applies from one module M_X to another when in M_1 a fully interpretable object for M_2 is assembled. Within FL,

however, even though conceptually our definition is dynamic, against Grohmann's *static* conception, the empirical results are surprisingly similar in *most* cases. However, there are a number of differences that have to be taken into account. In my opinion, the three domains are defined as *collapse areas*, providing *locative, eventive* (either agentive or not) or *temporal-aspectual-modal* information to the semantic interface. Even though the inclusion of Mod within the same domain as T and Asp may be objected to as Mod is rather related with so-called *discourse-driven* processes (and, thus, with the Ω-domain), I argue that those apparent processes are really post-syntactic, related with the concepts of *explicature* and *implicature*. Rizzi's movements to the Left Periphery nodes TopP and FocP in order to represent shared and new information, to put an example, cannot be feature-motivated in our framework. Let us put aside the question whether those movements take place in the syntax proper or not. If these last-resort operations have an effect on the interface such as the generation of an implicature is something irrelevant for the study of syntax alone, but essential to the study of the syntax-semantics interface, as inferential contents like entailments and presuppositions are syntactically determined. Mod is a node that participates actively (and sometimes independently of Asp and T) in the definition of the reference of nominal constructions as I have shown in Krivochen (2010a), and therefore, makes an essential contribution to the explicature in terms of referent assignment.

In Grohmann's model, *movement* is restricted in two ways: *locality* and *antilocality*. **Locality** establishes that an element must move cyclically, either inter- or intra-clausally, following these rules:

(47) *Intra-Clausal Movement Generalization*

$[_{\beta\Delta}$ XP … $[_{\alpha\Delta}$ … ~~XP~~ …]], where β >> α

Intra-clausal movement takes place between immediately dominating domains within a clause. Cyclicity means thus an intermediate landing point in each PD.

(48) *Inter-Clausal Movement Generalization*

$[_{\alpha\Delta}$ XP … ‡ … $[_{\alpha\Delta}$ … ~~XP~~ …]], where ‡ = clause boundary

Inter-clausal movement takes place between identical domains in different clauses. Cyclicity means thus landing sites in positions in identical PDs across clause boundaries, i.e., once a Ω-domain is complete.

Anti-locality establishes a restriction with respect to the occurrences of a SO within a PD, expressed in the *Condition on Domain Exclusivity* (Grohmann, 2003:107):

> *"For a given Prolific Domain ΠΔ, an object O in the phrase-marker must receive an exclusive interpretation at the interfaces, unless duplicity of O yields a drastic effect on the output of that ΠΔ."*

This means that there can only be a single occurrence of a given object within a PD (either the highest or the lowest, there is no restriction to it), unless Copy-Spell Out rule applies afterwards thus yielding the aforementioned "drastic effect on the output" (in PF terms). Therefore, movement *within* a PD is heavily restricted.

Even though I agree with Grohmann's generalizations and CDE, because they are both theoretically elegant and empirically adequate, I have already presented out attempt to derive both Locality and Anti-Locality from interface conditions, making them *principled* and

conceptually necessary, and not rules that empirically apply but whose *justification* (in our technical sense) is obscure. Let us summarize the different possibilities for Spell-Out timing I have sketched so far:

Spell-Out applies for a term K:

a) At the merger of the *closest phase head* proceeding in a bottom-up fashion (PIC_1)
b) At the merger of the *next phase head* proceeding in a bottom-up fashion (PIC_2)
c) At the completion of the *Prolific Domain* to which K belongs (Grohmann, 2003, 2004)
d) Whenever a *fully interpretable object* for the interfaces (containing K) is assembled by means of free, unbounded Merge (Krivochen, 2010b, 2011b).
e) When the *Command Unit* in which K is merged is completed (Uriagereka, 1999, 2002).
f) Within a *transformational cycle* affecting K (Epstein & Seeley, 2002).

Notice that our definition is the only one that requires no additional notions or stipulations, it only appeals to the concept of interface conditions, which are conceptually necessary in a massively modular mind-brain. This is what I understand as *radical minimalism*.

5. Parasitic Gaps revisited:

It is commonly assumed in the literature that there are two types of gaps (i.e., base positions for displacement):

a) True Gaps: originated by (A-A') Movement of elements in order to satisfy some featural requirement of a higher functional head. These gaps are independently motivated and stand on their own, without the requirement of being licensed by anything other than a specific [EF] / [EPP] feature.

b) Parasitic Gaps: licensed by a TG in a syntactic object, within which the PG and the TG are coindexed. We will expand on PGs, as they are the focus of the present section.

Some commonly assumed properties of PGs are the following:

- PG are licensed in S-Structure (i.e., after the application of a transformational rule)
- Antecedent in an A'-position in the matrix clause: PG cannot be licensed by A-chains
- PG are always DPs
- PG cannot be licensed by "non-referential NPs" (Nunes, 2004)
- TG *cannot* c-command PG, as PGs appear in adjuncts
- The adjunct must be non-finite (most commonly, infinitival)
- PG chain and TG chain constitute two different chains, in order not to violate the CED (see Nunes, 2004 for a different view on the matter).

In this section, we will see some instances of gaps that cannot be classified as either TG or PG according to traditional criteria, which should be revisited in order to gain descriptive and

explanatory adequacy. This, we will argue, is achieved via the *External Token Merge theory of displacement* we have described in Chapter 2. Let us review a classic example of PG in English:

(49) Which paper$_i$ did you file *TG*$_i$ without reading *PG*$_i$?

We see that all the characteristics are present here. There is A'-movement to Spec-CP and a PG licensed by this A'-dependency in an adjunct (as PGs are in adjuncts, regardless their linear position, they never c-command the TG). In this case, Nunes' (2004) *Sideward Movement* or the more traditional explanations with an empty *Op*$_{[Wh-]}$ in Spec-CP in the non-finite adjunct, and a *trace* in the PG position apply. The operator, in due time, is (somehow) coindexed with the Wh-element in the matrix clause, which allows the interpretative component to establish a dependency. This said; let us analyze the following structures in Spanish and English:

(50) Quedó media botella de cerveza sin tomar *t* (TG/PG?)
There-is half a bottle of beer without to-drink t

(51) a. Tengo todavía muchas fotocopias por leer *t* (TG/PG?)
Have$_{1Sg}$ yet many photocopies left-to-read t
b. I have many photocopies left to read *t.*

We see that in these examples there are so-called "extractions" from an adjunct PP or non-finite clause (which clearly violates the CED) and there is no other gap (either A- or A') that can be said to legitimize the dependency, in traditional Minimalist terms. Moreover, there is

no chain at all to license the gap, in case it is taken to be parasitic. The distinction TG/PG, then, must be deeply revisited.

The question we must ask ourselves now is: "does TG/PG distinction make any sense"? Let us revise what we have said so far:

A "true" gap is where we expect "regular extraction", whereas a "p-gap" is a structural place where extraction is usually not possible without the presence of a licensing TG. Categorially, nevertheless, they both behave like Wh-traces, following the ECP. This proposal (Putnam, p.c.) has a disadvantage, however: one still has traces and properties assigned to them without counting, of course, the mere concept of "extraction", i.e., the vision of *movement* as *displacement* against which we have argued. The provisional conclusion is that, rather than conditions over extraction, we should look for conditions over the establishments of dependencies at the interfaces, defining "dependency" as follows (Krivochen, 2012a):

"***Definition 24:*** *if α and β are interface-associated via their coordinates in W, there exists a Dependency between α and β.*

Definition 25: *if* $\alpha \in W_X \wedge \beta \in W_Y$ *and either* $W_X \equiv W_Y$ *or* $W_X \not\equiv W_Y$ *and α and β are defined by the same n-plet of coordinates in their respective isodimensional Ws, α and β are bound and the dependency is called co-referentiality.*

Definition 26: *Reference is location of a symbolic object in the conceptual multi-dimensional space via LI's coordinates, which are interpretable by the conceptual system C-I.*

Definition 27: *Dependencies are read off at IL, not at W.*"

Definitions 25 and 26 are essential to understand the semantics of the so-called PG structures, since the TG-trace and the PG have crucially the same referent, which is a problem for traditional approaches if one follows the claim that reference is established via c-command of *Op.* over *variable* (e.g., proper name-common DP).

Our system, thus, *Merges-α* freely and *dependencies* are established at the interfaces, where interpretation procedures apply. In our case, the establishment of a dependency has to do with the relation the interpretative component can establish between the set of coordinates that define the position of α and β in isodimensional workspaces. As far as $C_{(HL)}$ is concerned, there are objects sharing format that are merged following CP and DFI, both "third-factor" interface requirements. In this sense, both PG and TG are derived possibly in parallel by a simple mechanism, which can be made explicit as follows:

(1) Merge a *token-α* in the object position within the {P} structure in W_X

(2) Merge a *token-β* in the matrix structure in W_Y

(3) C-I Interface establishes a dependency between α and β *iff definitions 24-27* obtain.

With this framework, which unifies all gap-like phenomena, we can better explain multiple problems, like the so-called "resumptive pronouns" and *weak-strong* island violation. Let us analyze the following examples (taken from Alexopoulou & Keller, 2003: 3):

(52) **No island violation**

a. Who will we fire *t*/him?

b. Who does Mary claim that we will fire *t*/him?

c. Who does Jane think that Mary claims that we will fire *t*/him?

(53) **Weak island violation**

a. Who does Mary wonder whether we will fire *t*/him?

b. Who does Jane think that Mary wonders whether we will fire *t*/him?

(54) **Strong island violation**

a. Who does Mary meet the people that will fire *t*/him?

b. Who does Jane think that Mary meets the people that will fire *t*/him?

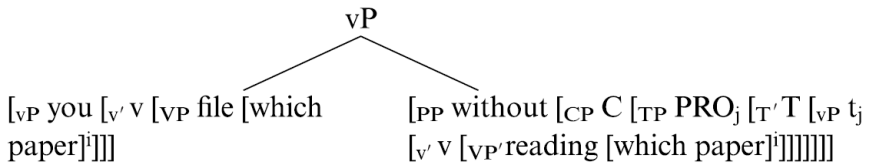
In our terms, all traces are eliminated, and resumptive pronouns are taken to be Spell-Out *tokens* of the same D-*type*, *á la* Grohmann's *Copy Spell Out*. Now, why should we bother in spelling those out, since we have already introduced an *Anti-Spell-Out* desideratum? Because *reference* chains require what we will call, in analogy to information theory, *signal reinforcement*: the origin of the "signal" is the first occurrence of a D-*token* in the linear order, and the *intensity* of that signal depends on the informational "weight" of the aforementioned {D}. In relevance-theoretic terms, the defining factors are *prominence*, *internal complexity* and *structure* of the ostensive stimulus. A {{D, √}, {Foc}} structure (i.e., a D with an embedded relative clause) is "heavier" (even in traditional terms) than a simple {D, √} structure, and heavier {D} are more difficultly forgotten by the parser (see also our discussion of passives and the licensing of [by-{D}] agents). Thus, the informationally heavier the {D} (and the more complex syntactically), the longer it will stay in the active working memory and the less the need for an immediate *signal reinforcer*, by which we mean simply *another Spelled-Out {D}-token*.

The advantages of our analysis over, for example, that of Nunes (2004), are quite straightforward: let us take an example like (55):

(55) Which paper did you file without reading?

Nunes' derivation of that parasitic gap supports our multiple workspace theory, since the complex object in (56) could not have been formed by *monotonic Merge* (Uriagereka, 1999):

(56)



Nunes says:

"*As the derivation proceeds, other lexical items are pulled out from* [a Numeration] *N' in (24a) and merge with K and M* [terms formed by Merge from N and N'] *in (25), yielding the objects P and Q in (26).*

(26) a. P = [$_{PP}$ without [$_{CP}$ C [$_{TP}$ PRO$_j$ [$_{T'}$ T [$_{vP}$ t$_j$ [$_{v'}$ v [$_{VP}$ reading [which paper]$_i$]]]]]]]

b. Q = [$_{vP}$ you [$_{v'}$ v [$_{VP}$ file [which paper]$_i$]]]

(27) [our (56)] *represents the next derivational step, where P adjoins to Q* […]. *In (27), no chain formation between the nondistinct copies of which paper can take place, since they are not in a c-command relation (...)*" (2004: 99)

We can see that his analysis needs to resort to parallel derivational workspaces, just as our analysis requires, but with the added complication of a trigger for *sidewards movement*, establishing either inter-tree dependencies or determining the derivational point in which a structure is non-monotonically merged. Moreover, the approach is grounded on c-command requirements, assuming an LCA framework and the mechanism of *chain reduction* Nunes proposes (2004: 27). So, from the four steps Nunes requires for *sidewards movement* to be

licensed in the syntax, we only leave one in the syntactic working area and the chain formation mechanism is made explicit in our definitions (24) and following:

(57) Nunes' (2004: 89) Copy + Merge theory of movement:

a. Copy

b. Merge

c. Form Chain

d. Chain Reduction

My *token*-Merge theory of displacement allows us to dispense with "Copy", since all we have to resort to is External Merge. Once a syntactic object is transferred, the relation between coordinates of terms results in a dependency, as we have defined them. A derivation of a PG, in our model, would replace Nunes' (26) with (58), more in line with Hale & Keyser's (2002) proposal on conflation:

(58) A. {you {*cause* {*event* [GO] {which paper {[TO] {file}}}}}

B. {Neg {you {*cause* {*event* √read {which paper}}}}

Notice that our steps (1), (2) and (3), made explicit above, unify the derivation of gaps without adding a further operation to the computational system outside External Token Merge. A {D, √} structure is assembled via monotonic Merge in W_X, and another token of that type enters the derivation in W_Y when a licensing node requires so to generate a drastic interface effect. Technically speaking, we have two distinct objects in $C_{(HL)}$, the dependency is only established at the semantic interface, and the choice of materialization copy depends on Spell-Out patterns in a language L, following a minimal effort desideratum: *Spell Out the*

minimal amount of tokens of each type needed to satisfy interface requirements, assuming semantics both drive and evaluate the derivation globally and locally.

6. Conclusion:

In this paper we tried to analyze Movement in three levels: *description*, *explanation* and *justification*. Whether we have been successful or not, is still to be proven. In this conclusion, we will sum up the main ideas of the paper:

i) All operations are interface-driven, and *every derivational step* must be justified in terms of generating a legible representation (*Dynamic Full Interpretation*).

ii) Movement, in terms of "interpreting something in a different place from which it is phonetically realized" (as opposed to movement interpreted literally as "displacement of a SO") is triggered by interface reasons, like *themehood* or *presuppositions*, what we have called "drastic interface effects". All the same, we state the pre-eminence of the semantic component over the phonological component. The generative engine, needs semantics to have something to manipulate[13], but can perfectly dispense with phonology, as externalization is not a *sine qua non* condition for syntactic manipulation of symbolic objects.

iii) The strong version of RM posits that:

Agree *is impossible if one does not have features (i.e., valued dimensions, see Uriagereka's comments to Chomsky, 1999), and it should be replaced by interface*

[13] As Immanuel Kant points out, in his Critique of Pure Reason, "*Gedanken ohne Inhalt sind leer, Anschauungen ohne Begriffe sind blind*".

reading algorithms. Dependencies must be read from the structure without adding stipulations.

Nothing "moves" in a literal sense, LA contains ***types*** *and the number of needed* ***tokens*** *is determined by the* ***Conservation Principle***.

References:

Acedo-Matellán, Víctor & Jaume Mateu (2010) From satellite-framed Latin to verb-framed Romance: A syntactic approach. Universitat Autónoma de Barcelona. Departament de Filología Catalana.

Adger, D. (2008) A Minimalist Theory of Feature Structure. lingBuzz/000583

(2011) Labels and Structures. lingBuzz /001252

Adger, D. & P. Sevenoius (2010) Features in Minimalist Syntax. lingBuzz/000825

Alexopoulou, T. & F. Keller (2003) Linguistic Complexity, Locality and Resumption. In *WCCFL 22 Proceedings*, ed. G. Garding and M. Tsujimura. Somerville, MA: Cascadilla Press.

Boeckx, C. (2006) *Linguistic Minimalism: Origins, Concepts, Methods and Aims*. OUP.

(2008) Treelets, not Trees. *BCGL 3 — Trees and Beyond*. May 21–23, 2008

(2009) The Nature of Merge. Consequences for Language, Mind and Biology. In Piatelli Palmarini et. al. *Of Minds and Language*. Oxford, OUP. 44-57.

(2010) Defeating Lexicocentrism. lingBuzz/001130

Carruthers, P. (2005) The Case for Massively Modular Models of Mind. In R.Stainton (ed.), *Contemporary Debates in Cognitive Science*. Oxford, Blackwell, 2006. 3-21.

Carston, R. (1998) The Relationship between Generative Grammar and (Relevance-theoretic) Pragmatics. Ms. University College, London.

Chomsky, N.:(1981) *Lectures on Government and Binding*, Dordrecht, Foris.

(1986) *Barriers*. Cambridge, Mass. MIT press.

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

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

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

(2001) Beyond Explanatory Adequacy. Ms. MIT.

(2005a) *On Phases*. Ms. MIT.

(2005b) Three Factors in Language Design. *Linguistic Inquiry* Vol. 36 n° 1. MIT press. 1-22.

(2007) *Approaching UG from below.* Ms. MIT.

(2009) Opening Remarks. In Piatelli Palmarini et. al. *Of Minds and Language*. Oxford, OUP.

Collins, C. (2002) Eliminating Labels. In *Derivation and Explanation in the Minimalist Program*, ed. S.D. Epstein & T.D. Seeley. Oxford: Blackwell. 42-62.

De Belder, M. & Van Craenenbroeck (2011) *How to Merge a Root*. LingBuzz 001226.

De Lancey, S. (2001) *Lectures on Functional Syntax.* University of Oregon.

Embick, D. & R. Noyer (2004) *Distributed Morphology and the Syntax-Morphology Interface*. Draft: 25/10/2004

Epstein, S. (1999) Un-Principled Syntax and the Derivation of Syntactic Relations. In *Working Minimalism*, S. Epstein, and N. Hornstein, eds. M.I.T. Press, Cambridge, MA.

Epstein, S. & T. Seeley (2000) *SPEC-ifying the GF Subject: Eliminating A-Chains and the EPP within a Derivational Model*. University of Michigan.

(2002) Rule Applications as Cycles in a Level Free Syntax. In *Derivation and Explanation in the Minimalist Program*, ed. S.D. Epstein & T.D. Seeley, 65-89. Oxford: Blackwell.

Escandell Vidal, M.V. (2006) *Introducción a la Pragmática*. Barcelona, Ariel.

Escandell Vidal, M.V. & Leonetti, M. (2000), Categorías conceptuales y semántica procedimental. In *Cien años de investigación semántica: de Michél Bréal a la actualidad* Tomo I, Madrid, Ediciones clásicas, 363-378.

Fábregas, A. (2005) *La definición de la categoría gramatical en una morfología orientada sintácticamente: nombres y adjetivos*. PhD Thesis, Universidad Autónoma de Madrid.

Gallego, A. (2007) *Phase Theory and Parametric Variation*. PhD Thesis, Universitat Autónoma de Barcelona.

(2010) *Phase Theory*. Amsterdam, John Benjamins.

Grice, H. P. (1975) Lógica y conversación, in Valdés Villanueva, Luis Ml. (Comp.) (1991) *En búsqueda del significado.* Madrid, Tecnos.

Grohmann, K. K. (2003) "Prolific Domains. On the Anti-Locality of Movement Dependencies". Amsterdam: John Benjamins.

(2004) Prolific Domains in the Computational System. *In Actas de JEL 2004: Domains*. Nantes: AAI, 211-216.

Hale, K. & J. Keyser (1993) On Argument Structure and the Lexical Expression of Syntactic Relations. In Hale & Keyser (1993) *The View from Building 20*. 53-110. Cambridge, Mass. MIT Press.

(1997a). On the Complex Nature of Simple Predicators. In A. Alsina, J. Bresnan & P. Sells (eds.), *Complex Predicates*. Stanford: CSLI Publications. 29-65.

(2002) *Conflation*. Ms. MIT.

Harley, H. & Noyer, R (1999) State-of-the-Article: Distributed Morphology. In *GLOT* 4.4. April, 1999. 6-9.

Hauser, Chomsky & Finch (2002) The faculty of Language: What is It, Who Has It and How did It Evolve? In *Science*, 298. November, 2002. 1569-1669

Heisenberg, W. (1999) *Physics and Philosophy*, New York: Prometheus Books

Hendrick, R. (2003) *Minimalist Syntax*. Oxford, Blackwell.

Hornstein (2000) *Move! A Minimalist Theory of Construal*. Oxford, Blackwell.

(2003) On Control. In R. Hendrick (ed.). *Minimalist Syntax*. Oxford: Blackwell, p. 6-81.

Jackendoff, R.(1983) *Semantics and Cognition*. Cambridge, MIT Press.

(1990) *Semantic Structures*. Cambridge, MIT Press.

(1997) *The Architecture of the Language Faculty*. Cambridge, Mass. MIT Press.

(2002) *Foundations of Language*. OUP.

Jackendoff, R. & F. Lerdahl (2004) The Capacity for Music: What Is It, and What´s Special About It? In *Cognition*, 2004. 33-72

Katz, J. & D. Pesetsky (2011) The Identity Thesis between Language and Music. Downloaded from http://web.mit.edu/linguistics/people/faculty/pesetsky/KP_paper_music_v2 5/11/11

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

Krivochen, D. (2010a) Referencialidad y Definitud en D. Un análisis desde la convergencia entre el Programa Minimalista y la Teoría de la Relevancia. UNLP / Universidad Nacional de Rosario. lingBuzz/001154

(2010b) Algunas notas sobre fases. UNLP. Conference presented at "I Jornadas de Jóvenes Lingüistas", UBA, March, 2011.

(2011a) A New Perspective on Raising Verbs. lingBuzz/001172 Submitted.

(2011b) An Introduction to Radical Minimalism I: on Merge and Agree (and related issues). In *IBERIA Vol 3 n° 2*. Pp. 20-62.

(2011c) The Quantum Human Computer Hypothesis and Radical Minimalism: A Brief Introduction to Quantum Linguistics. Published in *International Journal of Linguistic Studies* Vol. 5 n° 4, pp. 87-108.

(2012a) Towards a Geometrical Syntax: a Mathematical Formalization of Radical Minimalism. Ms. Universität Potsdam. Submitted. lingBuzz/001444

(2012b) *The Syntax and Semantics of the Nominal Construction*. Potsdam Linguistic Investigations 8. Frankfurt am Main, Peter Lang.

Krivochen, D. & P. Kosta (2011) Some (Radically Mimimalist) Thoughts on Merge. Presented at the Conference *Minimalist Program, Quo Vadis?* October 3-6, Potsdam University.

Krivochen, D. & P. Kosta (forthcoming) *Eliminating Empty Categories: A Radically Minimalist view on their Ontology and Justification*. Frankfurt am Main, Peter Lang.

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

Martin, R. & J. Uriagereka (2011) On the Nature of Chains in Minimalism. Talk delivered at the conference *Minimalist Program: Quo Vadis?* Universität Potsdam, October 3-6, 2011.

Mateu Fontanals, J. (2000a) *Why Can´t We Wipe The Slate Clean? A Lexical Syntactic Approach to Resultative Constructions*. Universitat Autónoma de Barcelona. Departament de Filología Catalana.

(2000b) *Universals of Semantic Construal for Lexical Syntactic Relations*. Universitat Autónoma de Barcelona. Departament de Filología Catalana.

(2008) *On the l-syntax of directionality/resultativity*. Universitat Autónoma de Barcelona. Departament de Filología Catalana.

Moro, A. (2000). *Dynamic Antisymmetry*. *Movement as a Symmetry Breaking Phenomenon*, Cambridge (Mass.): MIT Press.

Putnam, M. (2011) *The Thing that Should not be: A-A'*. Ms. University of Tromsø.

Rizzi, L. (2004) Locality and Left Periphery, in A. Belletti, ed., *Structures and Beyond – The Cartography of Syntactic Structures, Vol 3*, 223-251. Oxford University Press.

(2006) On the Form of Chains: Criterial Positions and ECP Effects. In L. Cheng, N. Corver, eds, *Wh-Movement: Moving on*. Cambridge, Mass. MIT Press.

Schrödinger, E. (1935) The Present Situation in Quantum Mechanics. in *Proceedings of the American Philosophical Society*, 124.

Sperber, D. & D. Wilson (1986a) Sobre la definición de Relevancia. En Valdés Villanueva, Luis Ml. (Comp.) (1991) *En búsqueda del significado*. Madrid, Tecnos.

(1986b/1995) *Relevance: Communication and Cognition*. Oxford. Blackwell.

Strawson, P. (1950) On Referring. In *Mind* 59, pp. 320-344

Tegmark, M. (2007) The Mathematical Universe Hypothesis. In *Foundations of Physics*, 2007. http://arxiv.org/pdf/0704.0646v2.pdf

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

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

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

Wilson & Sperber (2003) Relevance Theory, in Laurence Horn y Gregory Ward eds., *The Handbook of Pragmatics*, Blackwell, Oxford, 2004, pp. 607-32

Diego Gabriel Krivochen
Universität Potsdam, Institut für Slavistik
Am Neuen Palais 10, Potsdam, Germany.
Diego.Krivochen@uni-potsdam.de