An introduction to Radical Minimalism:

On Merge and Transfer (and related issues)

Diego Krivochen, UNLP. e-mail: diegokrivochen@hotmail.com

1. Introduction:

On this paper we will try to outline the basic tenets of ***Radical Minimalism***, going deeper in the line of what we have said in previous papers (Krivochen, 2010a, b, c, d, 2011). We will assume most of that framework, as well as orthodox Minimalism, and take that as a point of departure for new inquiries. We will test Radical Minimalism by analyzing what we consider to be the one and only generative mechanism in the human mind: the operation *merge*. We will review some work done about it and then present our own proposal, trying to derive the conceptual necessity and the properties of *merge* from interface conditions and then discuss the problem of *labeling*. On our way, we will also take the question of interpretable and uninterpretable features and the problem of valuation (Chomsky, 1999), and see if we can dispense with them in order to build a radically minimalist theory. We will also put our framework to test by presenting our theory of *transfer*, or *derivation by phases* in radically minimalist terms, which will implicate positing only *principled* operations and deriving everything from the dynamics of the derivation in interaction with interface requirements (see Boeckx 2010). Our ultimate goal will be to integrate the study of language (along with the study of any other mental faculty) in the study of the *physical universe* (or "natural world"), as a ***physical system*** among many others with which language should optimally share characteristics, operations and principles that we take to be *universal*.

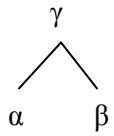
2. Development:

Part I: *Merge*

According to Chomsky (2007: 3, 5), "*In its most elementary form, a generative system is based on an operation that takes structures already formed and combines them into a new structure. Call it Merge.* (…) *Suppose X and Y are merged. Evidently, efficient computation will leave X and Y unchanged (the No-Tampering Condition NTC).*" Merge is restricted to two elements for theoretical and empirical reasons: two is the smallest non-trivial union of elements, which is the optimal option in a minimalist framework. Besides, positing strictly binary Merge is supported, among other arguments, by Kayne's *unambiguous paths* (which can be seen as a condition on phrase markers formation) and semantic interpretation (related to the problem of *labeling*, which we will

analyze deeply). Any generative theory is, then, in essence, a theory about Merge and its motivations and consequences. The type of system that Chomsky describes generates structures like the following:

i)



Note that Chomsky's remark is neutral with respect to the mental domain in which the operation applies, so that it is on us to restrict it to the Faculty of Language (FL from now on) or take the stronger view that *merge* is the one and only combinatory mechanism in the human mind, regardless of the module. As we have argued elsewhere (Krivochen, 2010b, d), we take the second, strong position. Any module that deals with structured discrete symbols must apply *merge* in order [1]to combine them, be them semantic primitives (see Mateu Fontanals, 2000a, b), sounds (Jackendoff and Lerdahl, 2004, Katz and Pesetsky, 2011) or lexical items, to give just a few examples. *Merge* is the minimalist name for Humboldt's remark that language is ***a*** system of discrete infinitude, to which we will add "but not the only one". The nature of α and β in our representation, then, can vary depending on the vertical faculty we are dealing with. The only thing we can say for sure is that this operation belongs to Chomsky's (2005b) *first factor*, genetic endowment, that is, it is not acquired but given. *Merge* is an operation that "comes free", which we will take in the sense that (a) it is computationally costless and (b) it cannot be reduced or decomposed. We will analyze this second point deeper in the following sections, when we review four of the most influential conceptions about *Merge* that are circulating nowadays, some of which try to decompose *merge* in more primitive operations.

2.1 Chomsky (1998, 2004, 2005a, 2007):

Given a *generative system* as characterized above, we need something to combine, and this something will be, for the time being, lexical items. However, this is not it, according to Chomsky. We also need a certain property in the units that lets them be combined, say, a feature conveying the syntactic instruction "I am mergeable". This property is called an *Edge Feature* (Chomsky, 2005a), and is present in every lexical item, excluding vocatives and interjections. Proposing this property (it is not a feature, really, since it is not a "valuable dimension", in Uriagereka's terms) amounts to trying to justify Merge syntactically, that is, Merge is a *feature motivated operation*, like anything else in $C_{(HL)}$. Of course, in a strict sense, if we take this view, Merge does not "come

[1] The continuous application of Merge is called "monotonic Merge". Merge's asymmetry comes from its intrinsically diachronic character (Epstein, 1999), *contra* Chomsky (1995), to whom its asymmetry was given by headedness. We will argue that, in consonance with Chomsky (2009), headedness is an *epiphenomenon*, interface-motivated.

free" at all, contrarily to Chomsky's claims, since we have to introduce a new element in the model, with the only purpose of justifying an operation. This position of motivating Merge has been also defended by Pesetsky and Torrego, who claim that there must be some kind of feature checking taking place, as *all operations should optimally be feature-driven*. The role of the EF is not clear nowadays, because it has been posited that this property, that allows Merge at the periphery (in consonance with the Extension Condition), is present in *phase heads* as a non-erasable element, thus allowing *multiple Spec-* positions that act as escape hatches and intermediate landing sites for successive-cyclic movement. The EF also overlaps with the old EPP feature especially in a *feature inheritance* (or *sharing*, as Gallego 2010 prefers to say) framework. Apparently, *phase heads* bear an EF, which can act as a double probe: in the case of C, for example, its EF, together with its Wh-feature, attract a Wh- marked constituent to the periphery of C, while it also attracts the EA to the Spec-TP position. Bong-Kim (2010) has suggested that, given the fact that V and T inherit features from their dominating-phase heads (*v** and C respectively), the EPP feature in T can be reduced to (or expressed in terms of) an inherited version of the EF in C, without clarifying, in any case, which exactly is the nature of this "inheritance" process or the EF itself.

GB's *Move α* has also been analyzed as an instance of Merge. If a derivation starts from a *lexical array*, and items are taken and merged in turn in the working area (External merge), "moving" something is merely re-merging a constituent, or, actually, a *copy* of that constituent, which can be seen as another occurrence of that particular SO. Movement has been renamed *Internal Merge* for that reason, a probe α seeks within its c-command domain a suitable goal β, features are matched, [u-F] are valued and then, if an EF in α requires so, β is *internally merged* in the periphery of α, leaving behind a copy which will not (normally) be Spelled Out.

Merge also brings the problem of *labeling*, how to signal headedness and account for endocentricity, since it seems to be an important feature of human language (we will return to this point later on). Chomsky attempted to solve it with a simple rule: he proposed that there were two kinds of *merge*, *set-merge* and *pair-merge* (Chomsky, 1998: 58). In the former, we are talking about adjunction, which is still a problem in minimalist theory, since no satisfactory theory has been yet proposed (that we know of). In those cases, if we *externally-set-merge* α to β, it is always β that projects. Chomsky has suggested (2004) that adjuncts are assembled in a parallel derivational space, and then introduced in the main tree by means of an old mechanism: a *generalized transformation*, which, simplifying, introduces a whole tree in another. Asymmetries between arguments and adjuncts are thus theoretically enhanced. In *pair-merge*, there is some "requirement" of α which is satisfied by its merger with β (say, argument structure), and it is α that projects. The labeling algorithm that Chomsky proposes can be summarized as follows (Chomsky, 2005):

a) In {H, α}, H an LI, H is the label

b) If α is internally merged to β, forming {α, β}, then the label of β is the label of {α, β}.

The problem is that this so-called algorithm is highly stipulative, and thus has to be replaced with an interface-driven procedure to make this labeling *principled*. As regards *pair merge*, we would need to encode the requirement relation between α and β in some form, most probably as a feature (or a subcategorization frame), an option against which we will argue, since we put forth (along the lines proposed by Boeckx and Gallego) that *Merge* is unbounded. Besides, our goal will be a *restrictivist theory*, with a strong presence of the interface conditions, not a *constructivist theory* like the one devised by Chomsky. The disadvantage presented by the second kind of theories is that, as their goal is to generate only convergent objects, they need a number of filters / *ad hoc* principles within the computational system itself, therefore preventing some mergers not because of interface conditions, but because some stipulated syntactic filter. Our model will aim to construct a *restrictivist theory*, in which Merge is free and the only "filters" for representations are the interface conditions. These conditions determine the set of what is known as "convergent derivations", a subset of the *possible* derivations. Although one may argue against restrictivism by saying that it is more costly than constructivism since all kinds of objects may reach the interface levels, this can be easily counter argued. We must think of derivations not only in *computational* terms but also in *biological* terms (that is, we have to take into account biological plausibility when positing an operation or a constraint on derivations / representations), as information is transmitted between connected neurons. Recurrent connections are rutinized (i.e., automated), and non-convergent derivations, although computationally possible (since, as we have said, Merge is unbounded and free), are ruled out by *interface conditions* at first and by *statistics* later on[2]. We must also take into account that, if we choose a Distributed Morphology framework with Late Insertion, we operate with the following principle[3]:

Morpheme formation constraint:

We cannot group features in a terminal node (i.e., morpheme) if there is no vocabulary item in the B List specified enough to be inserted in that node.

Of course, we have to distinguish yet *"uncoined" words* (any neologism, for example) from *impossible words* (a word that requires Spec-to-head incorporation, for example. See Uriagereka, 1998). Impossible words are banned because of syntactic principles, and therefore are not even formed. Uncoined words may be formed in the (l-)syntax if the intention of the speaker, already in the form of a Relational Semantic Structure (RSS, see Mateu, 2000a, b), requires so, and a vocabulary item may be created *ad hoc* following the regular morphological rules of the language in question to be inserted post-syntactically. For example, we can imagine a *location* verb

[2] An argument in favor of statistical learning within Generative Grammar is provided by Thornton & Tesan (2006).

[3] J.L. Stamboni, p.c. His view, however, differs from ours in a number of points.

[to table], but there is no way of forming [to wind] if [wind] is the external argument, as in [the wind blows], since an "illegal" movement would be taking place.

As a provisional conclusion, we may say that *the operation Merge may be blind, but the mind-brain is certainly not*. Interface conditions (which mediate in the communication between mental faculties) reject any malformed representation when it reaches the interface levels, since our theory is *restrictive*.

2.2 Hornstein & Pietroski (2009):

These authors take Merge to be a composite operation, made up of (a) *Concatenate* and (b) *Label*. The first operation merely puts ***two*** things together, be them simple or complex (without any implications for linearization purposes, apparently), and the second operation, applied to the results of the first, leaves that complex {α, β} as a single unit for the purposes of further operations (thus avoiding generation of syntactic objects -SO- like {α, β, γ}, which would not yield a linguistic expression), the label indicating which is the *head* of the new syntactic object, taking that every symbolic hierarchical structure is, in fact, endocentric. *Concatenate* operates only on "atoms" (whichever their nature), SOs that are already labeled and are therefore considered atomic for the effect of a new *concatenation*. The *label* operation must take the (categorial) features of one of the elements and project them so that the result of the operation is that the label is either α or β. Hornstein & Pietroski postulate that, apparently, *concatenate* is an ancient operation, available in animal cognition as well as human cognition, in their own words, *concatenation is an instruction to conjoin monadic concepts*, and nothing else. This can be seen as a deviation from the point of view exposed in Hauser, Chomsky and Fitch (2002), to whom recursion (the core characteristic of Merge) is an exclusive feature of FLN, and thus not available in non-human systems. Hornstein & Pietroski find a "*semantic correlate*" of *concatenate*, to account for compositional meaning, that is, the meaning of the whole depending on the meaning of the parts and the relations between them. We do not understand why it is necessary for them to distinguish these two instantiations of a combinatory operation, but our objection will be clear only when we have already exposed our own view.

In the same way they have decomposed (external) *merge*, they decompose *move* (i.e, internal merge). The whole sketch would look like this:

ii) Move = [Merge [Copy]]
Merge = [Label [Concatenate]]

The *label* operation is said to have roots in EST, resembling rewriting rules. In this framework, labeling is taken as "interpreting as…" Thus, simplifying on behalf of clarity,

iii) VP ⟶ V, NP

Amounts to saying "the node VP must be *interpreted as* composed by V and NP". The dynamics of *merge* would work the other way around, we first *concatenate* {V, NP} and then *label* the construction as a VP. However, we are not told how this labeling algorithm is working, that is, it is not made explicit. As a consequence, there is no principled reason why {V, NP} should be labeled VP (or simply V in a BPS framework) and not NP, apart from mere intuition or intra-theoretic stipulation.

2.3 Boeckx (2006, 2009, 2010):

Boeckx's proposal also involves decomposing *Merge* into operations that, separately considered, might be available in other mental domains than language, but together they are apparently unique to the Language Faculty. According to Boeckx, endocentricity in hierarchical structures is very language-specific (or, put in different words, part of FLN), and is made possible by two different operations:

a) Concatenate (or *Basic Grouping*)
b) Copy

These operations are presented by Boeckx as an instance of what Stephen Jay Gould called *Umbildung*, that is, recombination, using things already available in separate biological systems to form something new. These ancient operations taken separately can be said to be part of FLB as they are used in other faculties, but combined they characterize and define FLN. Contrary to what some have said, *concatenation* does not entail linear order, but simply "putting two things together" (whatever these things turn out to be, as long as they share "format"). The *copying* operation is a bit more complex, as it entails giving the newly formed structure the character of one of its components, or, what is equivalent, establishing which its head is. If we do this, we can use the whole structure (complex as it might be) as a unit for the purposes of new combination.
Combination is neither determined in any way by nor sensible to the nature of the elements that it combines, for if it were, then it could never be regarded as an operation available in multiple faculties. *Copying*, however, is of a more obscure nature, since it must have some access to the categorial features or equivalent (in the case of language) of the elements it manipulates in order to copy them and label the new structure. This can be implied from the fact that the result of merging X and Y is not a new structure Z, but either X or Y. In order to do this, the system must see something to copy; the first thing that comes to our mind would be *categorial* features, as we have already said. The results of these operations are *hierarchical endocentric binary-branched structures*, which are, in his view, unique to human language.

The nature of the combined elements is also discussed. Boeckx (2010) makes a distinction between *concepts* and *conceptual addresses*, which is parallel to that made in Krivochen (2010d) between *generic concepts* and *roots*. *Concepts* are not linguistic entities but more abstract elements which we take to be carry some sort of "meaning", and as such, they may be shared with other species as they are accessed by other faculties apart from FL, like the visual module (see discussion in Krivochen, 2010d). *Conceptual addresses*, on the other hand, are (pre-syntactic) linguistic entities which "point to" concepts (or convey instructions to "activate" concepts), and they are freely merged because they share a *common format*, not because there is a feature like Chomsky's EF that is driving merge. It would take a special stipulation to restrict Merge of conceptual addresses in the syntax, any restriction or constraint one would like to posit (in a theory like Boeckx's 2010 or our own) comes into play in the interface levels, post-syntactically. That is, we think, an important step towards a really minimalist theory of language.

Boeckx (2010) points out, and we agree, that feature bundles cannot be driving syntax, since these bundles are structures and that structure is (must be) syntactic in very much the road taken by Distributed Morphology (Halle & Marantz, 1993, Embick & Noyer, 2004, Panagiotidis, 2010), in which it is considered that roots and "f-morphemes" (to use the old DM term) are combined syntactically, with the same constraints that apply to an "s-syntactic" representation (like HMC, etc.). The elements the syntax manipulates should optimally be *atomic*, and Merge should be taken, in his view, as a free-triggered unbounded operation, *Merge α*. The whole argumentation of Boeckx's aims at "defeating lexicocentrism", that is, the presence of a pre-syntactic instance where fully-fledged lexical items are taken from. Contrarily to Chomsky's (2002, 2005b) claims, the "great leap forward" (that is, the qualitative evolutionary difference between humans and non-humans) would be *the emergence of conceptual addresses, not Merge itself*, since it is really a recombination of pre-existing processes.

2.4 Our view (Krivochen 2010b, d):

We have already presented our view on *Merge* on previous works, but here we will try to go deeper in our analysis, seeking the motivations for the mergers and trying to implement the whole machinery in a radically minimalist restrictivist theory. First of all, we have to address two questions:

a) Is Merge *feature-driven*?
b) Do we really have to decompose merge?

Our answer for both questions will be *no*. We will now address each in turn, analyzing what happens in $C_{(HL)}$ and then in the interface levels, to put some order in the architecture of the grammar.

Merge, as we take it, is a completely free operation that can apply as long as the objects to which it applies have the same format, motivated by interface conditions (this is, {α} is trivial in the interface levels, while {α, β} is not, as we have said[4]). In FL, we have lexical items[5], and we can say that they have the same format (be them "lexical categories" or "functional categories") since they share a nature, they are linguistic instantiations of elements that, *per se*, are not manipulable by $C_{(HL)}$. The only attribute of Merge would be putting things together, without any restriction by principle as regards the nature or number of objects, since it would be a stipulation. Is *binarity* an interface requirement, then? Yes and no. Binarity is the *simplest-non-trivial* combination of elements, and syntax (in the broad sense) is fundamentally economical: there is simply no point in applying Merge to {α}, thus generating {{α}}, if the latter object is in no way "more legible" than the former. If {α} is not interpretable, for some reason, there is no motive to believe that {{α}} will be interpretable, since it reduces to {α}. But there is also an interface requirement, related with the problem of labeling, that makes *binarity* a *principled* property of *linguistic* hierarchical structures (Chomsky, 2005b). Because of these asymmetries, in Krivochen (2010b) we have made the (purely descriptive) distinction between *syntax in the narrow sense* and *syntax in the broad sense*. The former refers to the recursive combinatory procedure of FL, what is usually referred to as Merge. The latter refers to the recursive combinatory procedure no matter which module we are talking about (C-I, the faculty of music, the mathematical capacity). Of course, our hypothesis is that *there is only one generative mechanism in the human mind, and that is Merge*, but for the sake of clarity we have found useful to make the terminological distinction presented above. Optimally, of course, the theory should dispense with this distinction and recognize that there is only one operation whose apparent variations correlate to differences in the characteristics of the objects it is manipulating and the interface conditions the generated representation must fulfill, a factor totally external to, and different from, Merge itself.

Our hypothesis (and here we start answering the second question) is also that ***Merge does not entail labeling***, in fact, ***syntax*** (in the narrow sense) ***can dispense with labels***. Labels are used to indicate which is the head of the relevant phrase, that is, which element will determine the properties of the fully-fledged phrase. However, we

[4] Boban Arsenijevic (p.c) claims that "*{{a}} is non-trivial in at least one faculty: the arithmetic capacity. Hence, output conditions can't be that bare to favor a binary merge*". However, our position is that if Merge is considered to be an operation and we assume also a *dynamic* version of Full Interpretation that states that *any derivational step must be interface-justified*, that is, the application of any opperation must lead to a legible object, to apply Merge to a single object is trivial in any faculty. If {a} is already legible in the relevant interface level, then why apply Merge in the first place? It would be computationally redundant, and therefore far from Minimalist. We maintain that binary Merge is the minimal-maximal non-trivial option. We therefore reject any proposal of *unitary Merge* on interface grounds.

[5] As a matter of fact, we have *roots* semantically defective and procedural features that make them manipulable by the computational system, but we will use the term "lexical items" for the time being. See *infra*.

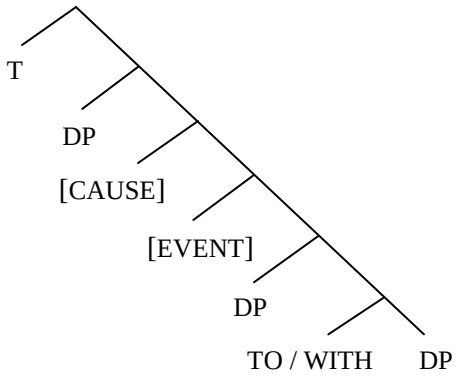
will argue, these properties are not relevant to syntax but to the C-I component in the construction of the explicature. Take, for example, (i):

i) Merge (love, Mary) = {love, Mary}

According to Gallego (2010: 15) in (i), the system sees no difference between [love] and [Mary], and he wants to formalize somehow that it is [love] that *selects* [Mary], and no the other way around, so labels are necessary in the syntax. However, our question would be the following: why would the system want to know that there is some "selection" relation if Merge is unbounded? If we have no s/c-selection anymore, no subcategorization frames in a "monolithic" lexicon, there is no way of representing this "selection" computationally. All the syntax can "see" is two elements with a common format (two lexical items in "traditional Minimalism", or, in Boeckx's terms, *conceptual addresses*) that can be put together with no restrictions apart from basic economy considerations. If they *can* be merged, then there is no principled reason why they should not be.

A hierarchical representation of a ditransitive structure in $C_{(HL)}$ could look like this (of course, it is not a complete tree, as we have comprised *Split TP* for the sake of clarity):

i)



Syntax only cares about combination or concatenation (in Boeckx's 2009 sense), because *there is nothing more than a generative mechanism, not an interpretative one*. However, the tree above would be illegible to the semantic component. In order to build an explicature, C-I needs an LF (in terms of Relevance Theory), a subspecified or incomplete representation of the propositional content with procedural instructions to *decode*, *disambiguate*, *assign referents* and *enrich semantically* those elements that need so.

Take the ditransitive structure above, for example[6]. If we merge D and P, that is all we do in the syntax, but in the semantic interface there are three options: either we do nothing, or we take the structure to be headed by D, or we take it to be headed by P. Labeling does not really matters to $C_{(HL)}$, because merge is unrestricted *there*, but it is a great way of signaling "headedness" either for the linguist (as a descriptive tool) and for the inferential component, the post-syntactic instance of C-I, but this does not amount to saying that *headedness* is actually a fundamental feature of language, as Boeckx and others have put forth. Chomsky (in discussion with Cedric Boeckx) argues that "*I don't think that headedness is a property of language. I think it is an epiphenomenon, and there is nothing simpler than Merge*". The "epiphenomenic" character of headedness is given by the fact that it is only required by a C-I requirement, but not by the narrow syntax. Returning to our ditransitive structure, if we "label" the resulting merger as D, all the procedural instructions of P are lost, no relation between *figure and ground* can be established because there is *neither a figure nor a ground*, as they are purely relational concepts. Of course, they do not exist in the narrow syntax, but the relation in nevertheless there, and it is P that establishes it. The same happens when we merge V, which introduces in the derivation an [EVENT] dimension. We can interpret it as two things: an event that includes a spatial relation or an extended spatial relation which somehow includes an event. Interpretation seems to be determined by *scope*, and since the eventive head c-commands the spatial relation, it has scope over it. Interpreting {V, {PP}} as P would lead the derivation to a crash, as we assume that there is a one-to-one relation between "projections" (in the sense of "labels") and types of information encoded in heads. That is, we cannot have a *single projection* (let us say, for the sake of clarity, a PP) with *two* features of different nature, conveying different information: a P head conveying spatial information relating *figure* and *ground* in terms of *central* or *terminal* coincidence and a V head conveying eventive information in terms of *telic* or *atelic* events. Thus, so-called "maximal projections" can be taken as "informational domains" in terms of interface conditions. That is, when monotonic merge builds an object and that object conveys a certain type of information which the next merged element does not convey, the "projection" is closed in terms of labeling, but this occurs *only because of interface conditions*[7]. The whole picture would look like this (we use traditional X-bar labels for the sake of clarity):

- {DP, {P, DP}} = P (domain of spatial information)
- {V, {PP}} = V (domain of eventive information)
- {*v*, {VP}} = *v* projection (domain of causativity)

[6] A thorough analysis of these structures and a justification for labeling is given in Krivochen (2010b, d), and we will not repeat that here, but add new information.

[7] The domains are "contextually defined", and in that sense, they may remind Grohmann's (2003, 2004) "Prolific Domains", but the reader must bear in mind the multiple differences between both theories.

- {DP (EA), {*v*, {VP}}} = full *v* projection.

We have a notational problem with the last two points. We have put "(full) *v* projection", without any bar-level indication, because we agree with Chomsky (1994) in the invisibility (or, more accurately, irrelevancy) of intermediate projections[8]. The question is: is there any relevance in XP (X'') notation? We think there is not. The [X]P is just a notational form of indicating that the relevant domain is closed, that is, the next head conveys another type of information. It is a form of signaling the end of a "contextually defined" domain in C-I terms. Bear in mind that this "label identification" process takes place *post-syntactically*, $C_{(HL)}$ manipulates bare trees like (i) above, whose "parts" are transferred as soon as a fully interpretable object is formed in terms of interface conditions, to which, of course, $C_{(HL)}$ must have access. By these means, we can dispense not only with labels in the syntax, but we can also simplify the labeling for LF (which we will always use in the sense of Relevance Theory) formation. No matter how we indicate it, *v*P, $v_{[+\text{máx}]\,[-\text{mín}]}$ or just *cause*, this fully-fledged SO is interpreted as a *caused event that includes a spatial relation*, and we have deduced all this without resorting to stipulations. The indications EA and IA (External and Internal Arguments), which are frequently used in papers on linguistics, and that we have used as well, are likewise nothing more than a descriptive tool, as "internal" and "external" argument distinction makes no sense if there is nothing to be external of in the syntax (that is, if there are no VPs at all and headedness is an epiphenomenon, interface-motivated). EA / IA / Loc, etc. are descriptive terms for (mostly) semantic concepts defined in the construction of the explicature, based on the position that DPs assume in the syntactic configuration, as we have posited in Krivochen (2010c).

Gallego (2010: 20), when discussing *phase-level labeling*, argues against approaches like our own or Collins' (2002) by listing some phenomena whose explanation and proper characterization allegedly *requires* labels:

a) Displacement / Pied Piping
b) Islands
c) Ellipsis
d) Prosodic Domains
e) A-structure
f) Incorporation
g) Endocentricity

[8] Both Kayne (1994) and Chomsky (1994, 1995) argued against intermediate projections, so that instead of $[_{XP}$ Spec $[_{X'}$ $[_{X}$ X] Compl]] we would have $[_{XP}$ Spec $[_{XP}$ $[_{X}$ X] Compl]]. Both had different theoretical reasons and, for us, Kayne's argument is a bit more solid than Chomsky's, since it relies on LCA (assimilating Specs- with Adjuncts). Chomsky simply stipulates that $C_{(HL)}$ can only "see" XPs and X_0.

h) Adjuncts vs. Complements

Our task now will be to try to demonstrate that either these phenomena do not occur in the syntax but in the way to the external systems or that if they occur in the syntax, they can be explained without resorting to labels. Of course we will not probably give a convincing argument for each, but our intention is clear: *radical minimalism* seeks elimination of labels in the working area, and if we can eliminate some other problems along the way, even better. These points will surely set the agenda for future research in our radical minimalism.

a) *Displacement*: the dynamics of movement are yet unclear, despite the huge amount of work done about it, since we claim against a feature-driven Merge, it would be contradictory to say that *internal merge* is driven by the need to check some feature or so[9]. It is not even clear where it takes place, if in the narrow syntax, as it has been the traditional view or in the PF branch, as Moro (2000) claims. If the latter, we should ask ourselves if PF is *label-sensitive*, which would have a negative answer in our framework, since we have argued that, if labels exist at all, they are only of relevance to LF, as informational domains seem to be irrelevant for prosodic purposes. This point is one of the most conflictive, and certainly requires further investigation.

b) *Islands*: islandhood should be subsumed to *phasehood*, in an optimal theory. That is, we see no point in having two locality conditions, two differentiated restrictions upon movement and extraction. By unifying all "barrier-like" phenomena, we are in the track of a much simpler theory, especially if we take a *principled* definition of *phase*, like our own (see Krivochen 2010b and below) or Grohmann's (2003, 2004).

c) *Ellipsis*: we see no reason to resort to labels in an explanation of ellipsis, if we say that a *term* can be elided only if there is a higher phonological occurrence of a parallel term (that is, we are talking of two different tokens, but with the same phonological form). In those cases, we can apply a generic vocabulary insertion in the lowest term, via "massive" fusion (in DM terms), say, a pro-form [do]. Of course, there is the question of what can be elided and what cannot, the determination of the aforementioned terms. That determination is not done in the syntax, but depends on what can be interpreted in the explicature level and the availability in the B-List of a pro-form specified enough in

[9] In Krivochen (2010c, 2011) we argue that movement of the subject to Spec-TP can be explained in terms of *theme-rheme* dynamics (and, therefore, it would be driven by a requirement of the semantic component), instead of resorting to an EPP feature. This theory has apparently been explored by Chomsky in his MIT 1995 courses (according to Uriagereka, 1998), but he has not explicitly formulated it in any paper we know of.

terms of features to fill the gap, for example, [do] (related with the T node in RSS) cannot replace just a spatial domain (i.e., PP), but there must be some [EVENT] primitive in the elided structure. A label-based ellipsis would work for a constructivist theory, but that is not our aim. Even if we are mistaken in this point (and we probably are), the basic argument is always the same.

d) *Prosodic domains*: prosodic domains are seen in non-generativist theories (like David Brazil's) as tone units, delimited by the occurrence of all the features of intonation (pitch level, pitch movement and prominence). Although these theories fail to provide an explanation of the phenomena involved and are of no scientific interest since they are actually unrestricted theories (there is no possible "malformation", but every unexpected choice is regarded as an "exploitation of the system" according to the speaker's intentions), the door is certainly open for alternatives. Chomsky has said that CPs and *v**Ps are both of semantic and phonological relative independency, but that is not clear at all, either from the explanatory point of view (why those projections and not others?) or from the descriptive point of view (a tone unit can ignore an allegedly phase boundary, as in /$_{P}$ i WANT them to GO/, where the *v**P phase does not constitute a tone unit in its own right).

e) *Argument structure*: we have not much to add to what we have said in Krivochen (2010c), but we can raise the bet by saying that there is no such concept as A-structure, since it would constitute an unprincipled complication of the theory. Let us assume that the architecture outlined in Krivochen (2010b) is on the right track. If that is so, then we start from a RSS built with generic concepts obeying the intention of the relevant mind[10], and following the *conservation principle* (see below) all the way through the derivation in all relevant levels, instantiating the information in such a way that it can be read by the level in question, but *information cannot be lost in the process*. A radically minimalist theory, with a stronger componential character than Hale & Keyser's can thus dispense with the concept of argument structure, whose encoding terms (a feature? A structural template?) represent a serious problem. All we have is a bare phrase structure and an explicature constructed with the information provided by the syntax.

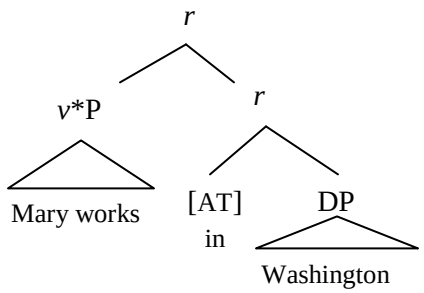
[10] This "intention" is something yet mysterious, since, as it is pre-linguistic, it should be formulated in non-linguistic terms, with which we would be admitting that there is thought without language, quite a strong claim. Bernárdez (1982) has used the term "global plan" to refer to this kind of intention, a purely pragmatic object after which semantics and syntax are shaped. We will not go deeper into this in this paper, but leave it to future research.

f) *Incorporation*: this process remains quite unclear, especially after Hale & Kayser's (2002) analysis of *conflation* as a distinct operation. According to their characterization, *incorporation* is a *movement* operation that depends on *government*, and thus incorporation from a Spec- to the immediate governing head is possible. Incorporation leaves a trace /copy, and the relations must obey the ECP (and so on). In a radically minimalist theory, there is no place for these kinds of processes, which rely on a stipulative relation (government, in this case). *Conflation*, on the other hand, is not a movement operation, but it is concomitant of Merge. It involves only the copying of the p-signature of the strict complement of a given head when this head is phonologically defective (i.e., an affix or null). However, even H&K resort to *labels* in their account, as Merge (according to them) creates labels given the fact that "*the label of a syntactic object X is the feature set [F,H], where [F,H] is the entire complement of phonological, morphological, syntactic and semantic features of H, the head of X"* (Hale & Kayser, 2002). Their *label identification* rule is not as arbitrary as Chomsky's, but it is equally dispensable in our model. If conflation is an operation on p-signatures (therefore, a PF operation, since a phonologically defective node represents no problem to syntax), there is no conceptual need to resort to labels at all.

g) *Endocentricity*: even though labeling is *the* way of signaling endocentricity, and here we agree with Gallego, we have said that syntax does not care about endocentricity (it is an "epiphenomenon", as Chomsky says), and labels are not relevant to syntax either (but to the C-I component), so there is no much left to be said at this point of the argumentation.

h) *Adjuncts vs. Complements*: we cannot say much about this, since we have not developed a theory of adjunction in our radical bare phrase structure. The first question to be addressed would be if the asymmetry actually exists, and, if it does, in what terms (for example, minimally required vs. not minimally required, which has been the traditional opposition). The theory of adjunction is intimately related to the concept of A-structure, so its current status is totally unclear to us. All we can say as a provisional answer to the problem is that all the information that we find in the explicature level (predicates, arguments, "adjuncts", etc) is already present from the very beginning of the process, due to the *conservation principle*. Information is neither created nor destroyed (or erased), but only transformed. Besides, if we are working with radically bare phrase structure, in which there are *no heads in the syntax*, there is no point in talking about Complement-adjunct distinction, since all we have is Free Merge, all other fine-graded differences being made in the explicature-building process. The "adjunction" structure with an unergative verb presented in Gallego (2010: 27), where the "main clause" is the *figure* and the "adjunct" is the *ground*, is an option that is worth considering, although it works only in a very limited number of cases, namely, those in which there is a spatial relation involved.

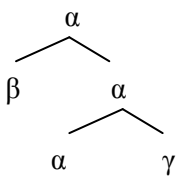
Logical relations, subspecified additive items (coordinations that really work as subordinations in the semantic level) and other types of adjuncts are not so easy to represent. Gallego's representation is as follows (adapted to RSS terms):



2.4.1 Monotonic Merge and Generalized Transformations:

At this point, once we have presented our theory of Merge, we must address a problem that was first expressed in minimalist terms by Uriagereka (1999), namely, *monotonic merge* and *parallel derivations*. We speak of monotonic merge when we apply the operation successively, forming what Uriagereka calls a "command unit", such as (i) below:

i)

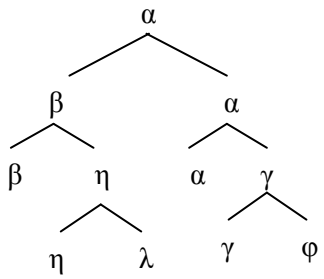


Here, we have no points of symmetry (i.e., symmetric c-command between two terminal nodes), so this structure can be linearized by LCA. However, for this system to work optimally, we must posit two conditions:

a) β must not be complex
b) γ (or the rightmost terminal in the structure) must be phonologically null, like a trace.

The second condition is the most difficult to solve, and no satisfactory answer has been given to it that we know of. Therefore, we will address (a), as it is not certain that (b) holds under Radical Minimalism[11] whereas (a) directly concerns our theory of Merge.

If β is complex, this is, {β, {β, η}}, then there would be symmetrical c-command between β and α in the following configuration:



There we have two command units, each assembled by *monotonic merge* simultaneously in separate derivational spaces (or sequentially in the same derivational space, an option that seems to have more biological plausibility, see Krivochen 2010d). So far, so good. The problem is that the system has to be able to merge *complex* units, whole trees[12]. We have said, following Boeckx (2010) that Merge is actually *Merge α*, a free unbounded operation that merges two objects that share format. Now, we would like to go deeper into that. We must distinguish two kinds of formats:

a) Ontological format
b) Structural format

Ontological format refers to the *nature* of the entities involved. For example, Merge can apply ("ergatively", as nobody / nothing "applies Merge" agentively) *conceptual addresses* (i.e., *roots*) because they are all linguistic instantiations of generic concepts. With *ontological format* we want to acknowledge the fact that a root and a generic concept cannot merge, for example. It is specially useful if we want to explain in simple terms why

[11] If there is no movement, then we must assume that a null node is inserted there, à la De Belder & Van Craenenbroeck (2011). We believe that such a theoretical resource only complicates the system substantively, and we will therefore dismiss this kind of solutions unless they prove not only inevitable but also optimal. At this point, we do not even assume LCA as correct, which is an obvious move if one takes into account that we are trying to de-construct minimalism in order to keep only what is conceptually necessary and what can be proven *principled* (i.e., interface-required).

[12] We are not making reference to Specs- and Adjuncts, since those notions have no place in our theory, but we are mainly thinking about Distributed Morphology and, most of all, Nanosyntax.

Merge cannot apply cross-modularly: ontological format is part of the legibility conditions of individual modules.

Structural format, on the other hand, refers to the way in which elements are organized. If what we have said so far is correct, then only binary-branched hierarchical structures are allowed in human mind. The arguments are conceptual rather than empirical, and we have already reviewed them: Merge optimally operates with the smallest non-trivial number of objects. Needless to say, given the fact that *ontological format is a necessary condition for Merge to apply* (principled because of interface conditions, whatever module we want to consider), the resultant structures will always consist on formally identical objects[13]. Taking all this into account, we must complete Boeckx's characterization of Merge and say that Merge is *a conceptually necessary operation that applies to the smallest number of elements sharing (ontological or structural) format*.

Given this characterization, we have a straightforward way of accounting for "generalized transformations", a recently resurrected operation from the very early stages of Generative Grammar. From our point of view, a generalized transformation is nothing more than *Merge applied to structurally identical objects*. The difference between "monotonic Merge" and "Generalized transformations" is the format that is relevant: *ontological format* is relevant to monotonic merge, whereas *structural format* is relevant to generalized transformations.

Let us review what has been said so far, in a very schematic way:

<u>Possibilities for Merge:</u>

1) Merge (α, β), $\alpha \neq \beta$ –but α and β share ontological or structural format- *Distinct binary Merge* (Boeckx, 2010)
2) Merge (α, β), $\alpha = \beta$ *Self Merge* (Adger, 2011)
3) Merge (α, β, γ…), $\alpha \neq \beta \neq \gamma$ *Unrestricted distinct Merge*

Let us consider the following:

j) *NS* Merge (α, β, γ) = {α, β, γ}

[13] This is not trivial if the reader is thinking about the possibility of merging structurally identical elements from different modules (which is obviously undesirable).

In the Narrow Syntax, everything would be fine, since Merge is blind and NS is ***not an interpretative component***. But, in the interface, problems would arise. Let us assume that α = √ and β and γ are procedural categories, say, D and T respectively.

k) *C*-I_2 Label {√, D, T} = ??

Having two procedural categories results in crash at the explicature level, there is no way of labeling a structure where two elements could "guide" the interpretation in different directions. The same happens if the numbers are changed, say, two roots and one procedural category: even if we think that one root may be "categorized" (which is not a viable option at all from our perspective), there would still be an uninterpretable element, namely, an uncategorized root, uninterpretable in LF because of semantic underspecification. Binary-distinct Merge, then, is interface-required, no special conditions imposed over Merge itself. In any of the cases, it must be said, the application of Merge involving Ø (e.g., {α, Ø}) is equal to Self-Merge for interface purposes, thus being rejected in our proposal.

A note on Collins (2002):

Collins also attempted to eliminate labels as they were conceived of and make use of label-free trees like our (i) above, doing away with Chomsky's labeling algorithm. No operation can make reference to "maximal projections" or "intermediate projections", since those terms only have sense within traditional X-bar theory. Collins' argument is that, even though lexical categories may have "categorial labels" like N, V, A, P, those categorial features *do not project*, and therefore, there are no VPs, NPs or anything like it. He focuses on four areas on which labels are apparently used:

- X-bar Theory
- Selection (in terms of subcategorization)
- Minimal Link Condition
- PF interface

According to him, all reference to labels can be dispensed with if we take into account four "syntactic relations" that can hold between lexical items, malformations being mainly a matter of Minimality:

- Theta (X, Y): X assigns a theta role to Y
- EPP (X, Y): X satisfies the EPP feature of Y

- Agree (X, Y): X matches Y, and Y values X
- Subcat (X, Y): X subcategorizes for Y

Even though we agree with his intention, we must say that Collins' solution does not seem right to us. His model can be proved wrong if we can demonstrate that those operations either do not exist or are not syntactic. Theta roles have been deeply analyzed in Krivochen (2010c) they are semantic functions that are read in the explicature level from the information that the syntax provides: case, structural position (both in RSS and "narrow syntactic" representation) and materialization. There is no such thing as theta role assignment in the syntax (or theta features, as Hornstein (2003) suggests), since they are *not relevant to syntactic computations*, but only to the construction of the explicature, a full propositional form.

As regards EPP, it has been eliminated in various ways. Some take it as an inherited version of C´s [EF] in T, but we think that that is quite a circular argument, since EF is as stipulative as the EPP. We prefer to stick to the idea that Spec-TP is a position that conveys only "*themeness*" in informative terms, given that structural position is meaningful in the explicature level. If this is so, then those nominal expressions that are *themes* will rise to that position, which is, as we argued in Krivochen (2011), *semantically transparent*, as no particular inference is related to it.

We have already dispensed with *Agree* as an operation, resorting instead to *licensing* (see Krivochen, 2010c and the discussion below). For us, dimensions are *licensed* in an element only if it is in a local relation with an appropriate procedural/functional head. That local relation can be described in terms of *phasehood* or *Minimal Configuration* (Rizzi, 2004), where H and XP are in an MC only if there is no "potential governor" between H and XP. Our theory of quantum dimensions attempts to simplify the whole agreement system, so we will leave the topic for the time being.

Subcategorization is the most difficult element to justify in a (radically) minimalist framework, since it has to be encoded in a feature which not only will add complexity to the representation but also impose a stipulative restriction on Merge. Merge would be an operation quite analogous to GB's *satisfy*, because theories with subcategorization frames are strongly *constructivist*, and their aim is to generate only legible structures by imposing constraints to the phrase formation operations.
We have thus analyzed different views on the labeling process, and we have tried to prove that there is no algorithm whatsoever to resort to in a radically minimalist framework, only the dynamics of unbounded Merge and the interface conditions.

Part II: *Transfer*

Transfer is the operation that takes a given piece of information and sends it to another module or faculty. The transfer to the A-P component is what used to be called *Spell-Out*, or *vocabulary insertion* in a DM framework. In a *massively modular* mind, *Transfer* is essential in the interaction between modules. In Krivochen (2010d) we have differentiated *transfer* from *transduction*, taking the former to be the transmission of legible information from one module to another, and the latter to be the translation of information coming from the *phenomenological world* to a given *vertical faculty*, in Fodor's terms. In that paper we also said that as soon as a fully interpretable object (in terms of the interface levels, whatever these are) is assembled in a given level, it is transferred, thus leading us to a non-stipulative definition of *phase*:

P is a ***phase in*** $\boldsymbol{L_X}$ *iff it is the* ***minimal term fully interpretable in*** $\boldsymbol{L_{X+1}}$

This definition is *dynamic* and presupposes a strongly componential architecture, in which the components can *look ahead* to the legibility conditions of the module to which they have to transfer information, in order to do so in a legible format. Information is taken to be carried in the form of features (Chomsky, 1995), be them phonological, semantic or syntactic. These features are, according to Uriagereka's definition, "valued dimensions" of the form [+/- D], being D a given *dimension* and two possible values, + and -. That is the canonical representation of features and the one we will use for the purposes of the present discussion[14]. Within the Minimalist framework, features were fundamentally divided (Chomsky, 1995) in *interpretable* and *uninterpretable* on the one hand and *valued* and *unvalued* (when entering the derivation) on the other. A feature was not (un-)interpretable *per se*, but depending on the category it was part of. The logics were the following: if a feature F makes a *semantic contribution* in a LI, it is *interpretable* in that LI. Thus, [φ-features] were uninterpretable in T, but interpretable in nominals since [person / number / gender] are (allegedly) semantically interpretable in DPs, and it is the verb that agrees with the noun, and not the other way around. Of course that reasoning is wrong, since (a) there is no definition of "semantic contribution", and (b) if by "semantic contribution" we take "contribution to the explicature" (*decoding* and *referent assignment*), we are in serious trouble since φ-features in nominals, for example, are sometimes tricky, like in the case of *pluralia tantum*. Realizing this mistake, Chomsky (1999) changes the angle and entertains the idea that *uninterpretable* features enter the derivation *unvalued* from the *lexicon*. In his view, syntax only cares about valuation, not interpretability. However, we also find problems with this proposal. To begin with, the stipulative character of the assignment of valued and unvalued features to different categories is maintained. Besides, it is a clearly syntacticocentric solution, with no attention drawn to the interface levels and interpretability in the external

[14] For details, see Adger (2008) and Adger & Sevenoius (2010).

systems. Unvalued features must be valued by *agree* between a *probe* and a *goal* in order to assure *convergence*. Once valued, the uninterpretable features are eliminated, by means of erasure. Interpretable information, on the other hand, is conserved, since it makes a "semantic contribution", whatever that means. Lasnik, Uriagereka & Boeckx (2005) borrow *Conservation Principle* from physics, and state the following law:

1st Conservation Law (Lasnik, Uriagereka & Boeckx's version):

All information in a syntactic derivation comes from the lexicon and interpretable lexical information cannot be destroyed.

The problem with this law is that it makes use of lexical information taken from a pre-syntactic and monolithic lexicon, which is the norm in orthodox Minimalism, but with which we will not work. However, keep in mind the spirit of the principle, since we will try to "recycle" it later. They go further away, positing a second conservation rule, applying to *interpretable structures*[15]:

2nd Conservation Law:

Interpretable structural units created in a syntactic derivation cannot be altered.

Given our definition of *phase,* we can understand what they mean by "interpretable structural units", in fact, one could easily replace that in our definition without any significant loss of meaning. But we will try to introduce a radical change in the way derivations are seen. For that purpose, we modify a bit the conservation laws to fit our presentational purposes:

First formulation of the Conservation Principle (our formulation):

a) *Interpretable information cannot be eliminated, it must go all-the-way through the derivation*

b) *Uninterpretable information is "viral" to the system and must thus be eliminated by Agree*

In this formulation we expect to be covering and summarizing the meaning that both Chomsky and Lasnik et. al. wanted to convey ("viral" theory of uninterpretable features is Uriagereka's).
After having presented all this information, there is still a basic question to be addressed: what is really the difference between interpretable and uninterpretable features? Or, to be a bit more radical: are there *uninterpretable features* at all? Notice that the difference between interpretable-uninterpretable, although subsumed to that between valued and unvalued, is still there at the very core of the theory, and is taken as a

[15] This principle is reminiscent of Edmond's *Structure Preserving Hypothesis*, but the mention of *interpretability* gives it a more "minimalist flavor".

primitive, with no explanation whatsoever. If syntax only cares about putting things together, why should it bother about valuation and so on? After all, *feature valuation is an operation that only makes sense taking convergence at the interface levels into account, but in the syntax proper* (or "narrow syntax") *it is perfectly superfluous*, since nothing "converges" or "crashes" in the syntax.

What we propose regarding all so-called "uninterpretable features" is that they ***do not exist at all***, especially considering the proposal made in Chomsky (1999) that *uninterpretability* is concomitant to *unvaluation*. That would be the strong (and optimal) thesis. Instead of a number of features (number that has increased over the years) which enter the derivation valued or unvalued depending on the category they compose, we have a minimal number of ***quantum dimensions*** conveying semantic (conceptual or procedural) meaning, which adopt one value or another in a *local relation* with a proper procedural head, namely, a *Minimal Configuration* (Rizzi, 2004). Let us analyze what we mean by *quantum dimensions*. Chomsky's feature valuation process needs:

a) A *probe*, an *unvalued* dimension in a (functional) head

b) A *goal*, the *same* dimension but *valued* in a c-commanded head[16]

We see a redundancy here, since we will depart from the claim that there is no need for the same dimension to be in two heads just for the sake of feature valuation, the arguments in favor of these operations become cyclical[17]. Our argument seeks to eliminate *Agree* both as a relation and as an operation. If the system is really "free", as Boeckx (2010) claims, then there cannot be any "Agree" operation as currently formulated since it would imply a complication for the theory and a restriction for the possible relations between elements (namely, relations are limited to those between *probe* and *goal*). A constructivist theory needs *Agree* in order to establish the "right" (i.e., convergent) relations between elements, but in our theory, based on *licensing*, such a restriction is regarded as stipulative.

[16] The relation is always head-head, even if we say that T matches features with a DP, it is really matching features with D, and all subsequent operations (e.g., movement) apply to the smallest term that assures convergence, namely, the whole DP.

[17] If we take an architecture like the one outlined in Krivochen (2010b), we would first have a "fully interpretable" RSS and then we would add (un-interpretable) features that would be later on valuated by Agree and erased from the computation (see Epstein & Seeley, 2002, for a discussion on this specific point). That is, for us, a redundant (and stipulative) computation, and plainly inadmissible in a Radically Minimalist theory.

Our claim here will be that language is part of the natural world, and as such, it is a system whose physical properties are the same as any other system. We will invoke here Heisemberg's (1927) *indeterminacy principle*, which can be better explained from an example:

Imagine we have an electron in a tridimensional space, and we want to know its location. In order to do so, we need to see it, projecting some kind of light on it. This light is projected in the form of a *photon*, a particle with mass. The "problem" is that when the photon crashes with the electron, there is a change in the original location, which remains unknown. That original location (we have taken this magnitude just for the sake of the example, but we could have also worked with speed or trajectory) is taken to be a "superposition" of all possible locations, expressed in the form of a "*wave function*" (in de Broglie's terms). Therefore, there will always be a magnitude whose real value will remain unknown to us. In this kind of physical systems, it is the ***observation*** that makes the relevant dimension *collapse* to one of the possible states[18]. *Indeterminacy is a natural characteristic of physical systems, and by no means an instrumental problem*, taking physical system in its technical sense, that is, any portion of the physical universe chosen for analysis. We take "physical universe" to be equivalent to "natural world", and we will use one or the other indistinctly. Magnitudes (or *dimensions*, to maintain a term more closely related to linguistics, since we are not dealing with measurable elements) are *not* necessarily binary; what is more, *in abstracto* they can comprise as many states as the system requires, which, as we will show later, leads us to a much simpler form of minimalism. ***We will express it by using this notation: for any dimension D, [D_X] expresses its quantum state.***

Let us suppose that we have a physical system which starts out in a state α, and changes, over some time, into state α'. Of course, it could have started out in any of many *different* states. So suppose it starts out in state β, and changes over the same considered time interval into state β'. We can schematically represent these two possible "trajectories" like this:

i) $\alpha \rightarrow \alpha'$

ii) $\beta \rightarrow \beta'$

Since α and β are possible states of the system, so is their arbitrary linear combination $\mathbf{a\alpha + b\beta}$. What ***Schrödinger's Equation (SE)*** tells us is that given that α and β would change in the ways just indicated, their linear combination must also change in the following way:

[18] See, for example, the well-known EPR (Einstein-Podolsky-Rosen) paradox, which inspired Schrödinger (1935) paper.

iii) **a**α **+ b**β ⟶ **a**α**' + b**β**'.**

The interesting fact about the above mentioned equations is that *they hold only if no "measurement" is taking place.*

If a "measurement" (say, mere observation) is taking place then we must consider an entirely different story about how the state of the system changes: during the measurement, the system S must "*collapse*" into a state that is certain to produce the observed result of the measurement. The hypothesis is exemplified by Schrödinger (1935) using the now famous "cat paradox", which deserves to be quoted in full-length:

> *"A cat is penned up in a steel chamber, along with the following device (which must be secured against direct interference by the cat): in a Geiger counter there is a tiny bit of radioactive substance, so small, that perhaps in the course of the hour one of the atoms decays, but also, with equal probability, perhaps none; if it happens, the counter tube discharges and through a relay releases a hammer which shatters a small flask of hydrocyanic acid. If one has left this entire system to itself for an hour, one would say that the cat still lives if meanwhile no atom has decayed. The ψ-function of the entire system would express this by having in it the living and dead cat (pardon the expression) mixed or smeared out in equal parts. It is typical of these cases that an indeterminacy originally restricted to the atomic domain becomes transformed into macroscopic indeterminacy, which can then be* ***resolved*** *by direct observation*". (p. 7-8. Highlighted in the original)

The question to be asked now is: *how do we apply this to language?*

Our answer will be the following: ***we will consider language to be a physical system, and therefore, if SE applies to any physical system, it must also apply to language***. Of course, we are not saying that language shares all of its features with other systems (since, as Boeckx correctly points out, we may have one or the other scattered in different systems), but it must be considered fundamentally as a physical system if we take seriously the idea that it is part of the natural world, as Chomsky has explicitly done along the years. We must say in this point that, despite what it may seem[19], there is no *reductionism* in treating FL as a physical system if we consider that a ***physical system*** is merely *the portion of the universe taken for analysis*. If we consider that universe to be the so-called "natural world", then, our thesis follows naturally. That is, we are not making a reduction of biology to physics, 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 Poppel & Embick, 2005 correctly point out), but we put forth that the *methodological level* has

[19] We thank Phoevos Panagiotidis (p.c.) for making this objection, and other valuable comments as well.

much to tell us, as we are all working with "parcels" of the same Universe that, we will try to show, are ***identical in a principled level of abstraction***.

The next step would be to put this theory in practice. Let us assume the following quantum dimension: [$Case_X$]. Following the idea presented in Krivochen (2010c), this dimension comprises three possible "outcomes": NOM sphere (φ), ACC sphere (θ) and DAT sphere (λ). All three are possible final states of the system, and therefore the linear combination must also be considered a legitimate state of the system. The dimension *in abstracto* could then be expressed as follows, using SE:

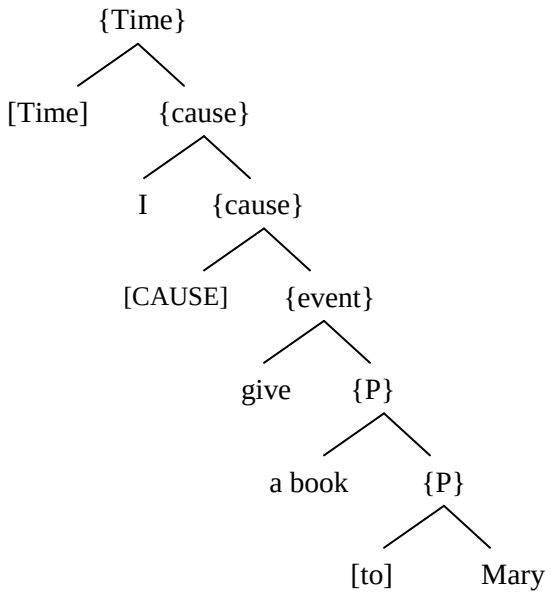
iv) Nφ + Aθ + Dλ

As we have said before, this only holds if no "measurement" takes place, in Schrödinger´s terms. We will not speak of "measurement", since Case is not a magnitude, but we will consider that the factor that makes the relevant dimension collapse is ***the merger of a functional / procedural head***. What we must take into account is that not only do we have DPs with [Case] and functional heads in interaction in the computational system, but the output (i.e, the resultant state) must also converge at the interface levels, so our problem is a bit more complicated. As usual, we will focus ourselves in the C-I component. What we want to do now is derive the relations P-DAT; *v*-ACC and T-NOM from interface conditions, apart from the argumentation we have made in Krivochen (2010c) in relation with θ-roles and Case, to which we refer the reader. Anything else would be stipulative, and that is something we cannot accept in Radical Minimalism.

Epistemologically, we have the advantage over theoretical physics that we have observable stretches of language where to test our hypothesis. So, let us take a ditransitive *Prepositional Indirect Object Construction* (PIOC):

v) I_{NOM} gave a $book_{ACC}$ to $Mary_{DAT}$

By looking at the same construction in other languages with rich casual morphology (like Latin, Sanskrit or Greek) it has been established that the DP [I] has Nominative Case, [the book] has Accusative Case and [to Mary] has Dative Case. But this is nothing more than a description, with no explanation as to why things are the way they are. We will not review the classic attempts of explanation, but go directly to our point. The (LF) structural configuration for (v) is the following (labeling according to interface conditions, as said above –and see Krivochen, 2010b, d), and omitting irrelevant positions / projections):

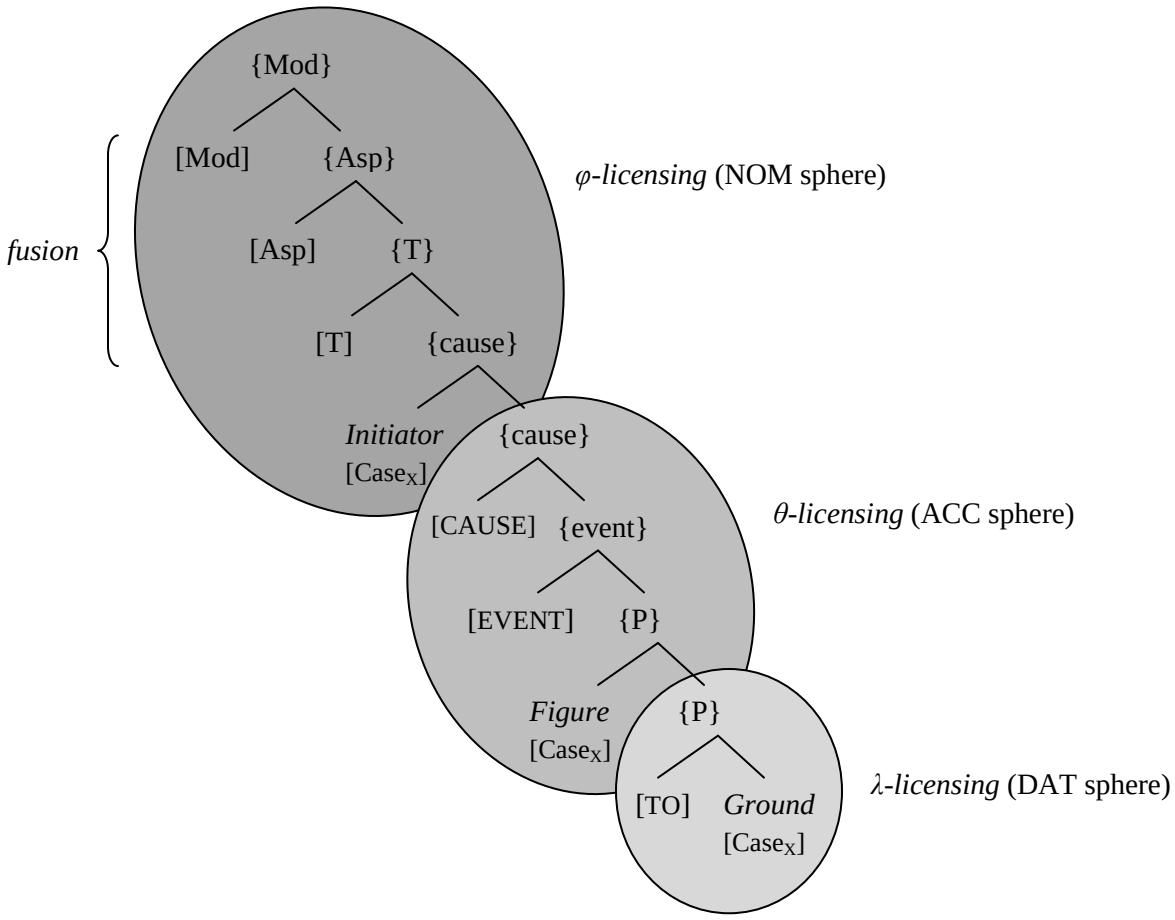


By hypothesis, each DP will bear a dimension [$Case_X$], whose value when entering the derivation will be a "ψ-function", or a complex vector [Nφ + Aθ + Dλ]. That is, the quantum dimension will comprise all three possible values, as all three are possible states of the Case system for a particular DP. However, [D_X] cannot be read by the interface levels, so the quantum dimension must "collapse" to one of the possible states. Here, it is not "measurement" but *Merge* that does the work. Let us take the {P} structure as an example. We have the merger of {$to_{[TO]}$, $Mary_{[Casex]}$} -and the subsequent "labeling" of the structure as {P} in LF for C-I interpretability at the explicature level-. At this point in the derivation, we are already in condition of collapsing the quantum dimension in [Mary], since we have a *procedural head* in a minimal configuration that can license that dimension in a c-commanded XP (adapting the idea from Rizzi, 2004). This minimal configuration is defined in our terms *within phase boundaries*, following the definition given above. Since the procedural information conveyed by P is essentially *locative*, the quantum case dimension on [Mary] will collapse to the locative sphere, i.e., Dative. The licensing takes place only if there is no closer functional / procedural head that can license the relevant feature, in order to respect Minimality. Just as we have taken SE and Heisemberg's indeterminacy principle from quantum mechanics, we can also draw another principle, this time from *field theory* (Lasnik, Uriagereka & Boeckx quote this principle from fluid mechanics, but it is the same principle, as we are ***always*** dealing with physical systems): ***Locality***. Just like a particle cannot influence any other than its surrounding particles, a procedural head has a certain area of influence where it can license features if necessary. Remember

that there is no checking / matching, so neither *greed* nor "enlightened self interest" are involved. It is interesting that influence can be indirect, that is, a particle α may not be able to influence particle γ since particle β is "in the middle", but by influencing β, α will have an effect on γ. That is what we call compositionality, in linguistics. We have to pay attention to the whole derivation (in *phase-level* terms) to understand, for example, the interpretation that a certain DP receives (see, for examples and analysis, Krivochen, 2010a): in that case, we have to take into account three procedural heads, T, Asp and Mod. Asp can influence a {D} structure in {cause}, but in conjunction with T, not on its own, since T is an intervenient head as they are part of the same informational domain. There is no blocking (because there is no *probing*), but "cumulative influence". However, it seems that this "cumulative influence" is possible only when the relevant nodes can be *fused* (in DM terms), and thus form a single morpheme prior to Vocabulary Insertion or, in terms more close to Grohmann's, when the nodes form part of the same domain. That is, after the configuration {Mod, {Asp, {Time, {cause}}}} is complete, compositionality does its work and the dimensions in each functional head combined determine the (most accessible) interpretation for the [Definiteness] dimension in D, then, in a language like Spanish, they can fuse morphologically and a single item is inserted in the terminal node. These dimensions, we have argued in Krivochen (2010a), are "valued" in a pre-syntactic instance, namely, C-I as their "collapsing" must be prior to that of the dimensions in D (since it is a *sine qua non* condition for {D} interpretation). By hypothesis, if FFCC are PPCC (as Leonetti & Escandell posit), and these encode instructions for the manipulation of conceptual information, these instructions depend on the intention of the speaker, a very basic and primitive representation whose nature, as we have already said, is yet unknown. However, even though P, and *v* are procedural heads, there is no cumulative influence but *blocking* since there is no possibility of *fusion*. Why is that? Because the {P} structure conveys a type of information, namely *spatial*, and that is totally different from and independent of *cause*, as they can appear separately (e.g., unaccusative structures have P but not *v*, and unergative structures have *v* but not P). What is more, the value of the dimension in *r* depends on the configuration the arguments (i.e., non-relational elements, see Mateu, 2000a, b) adopt in that structure, by which we mean that it is totally compositional. For example, in a verb like [shelve] (a book), a *central coincidence relation* would make the derivation crash at the explicature level, since a book cannot *have* a shelf, and we have seen (Krivochen, 2010d) that the node *r* can be clausal, as the primitive [WITH] is equivalent to the verb [HAVE] (equally primitive). In the case of T, Asp and Mod, there is no way they can appear separately or "scrambled", as *scope* determines interpretation: we cannot present an event as perfective or imperfective if we have not situated it in the time axis first, and by doing so, given it *reference*. What is true is that {Mod} may not appear in certain structures, like ECM, to allow influence from a functional head in the higher clause[20].

[20] See the argumentation in Krivochen (2011: 14 ff.)

The type of information that each domain carries in the form of dimensions is what makes the quantum dimensions in the arguments collapse, let us analyze this in a tree diagram (including labels only for clarity purposes):



Notice that the procedural / functional heads P, *v* and the *Split TP* signal *quantum feature collapse areas*, in consonance with Minimality. Optimally, all quantum dimensions on a certain element should collapse within the minimal collapse area (or informational domain, as we have said above). The three Case spheres, as we have already said in Krivochen (2010c), are in correlation with the Thematic spheres, following DeLancey's (2001) proposal. Projections are completely irrelevant, but we have included labels for the sake of clarity. Each circle determines the boundaries of a domain of information in terms of LF, the *locus* of collapsing. Collapsing cannot take place across boundaries for two reasons: first, because of the Earliness Principle: operations take place as soon as they can, and there is no reason to wait for another functional head to merge when we already have a head that can make the quantum dimension collapse. Second, because of Minimality: different types of

informational domains are like different perspectives for measurement, they make the dimension collapse to one state or the other. The difference with physics here is that there is a correlation between the information carried by the functional head and the state it licenses. Notice also that it would make no sense to talk about "VP" or anything like that, since VP presupposes a root already categorized, and we posit that category, like Case, is a quantum dimension and that it collapses in a local relation with a procedural head, as we will see below.

The collapsing of quantum dimensions can be taken as a possible justification for certain instances of *movement*: a quantum dimension in its ψ-state is not legible by the interface levels, as we have said. Let us suppose that we have a SO [X] with two quantum dimensions, $[P_X]$ and $[Q_X]$. We have said that there is a correlation between the information that a certain head carries and the dimensions it can license, not because of a syntactic constraint, but because of interface conditions: if a procedural node carries spatial information, and makes a definiteness dimension collapse, the result will be a crash in the explicature level, since that node -P, for the sake of clarity- is related with the Dative / Locative sphere (Krivochen, 2010c), but not with definiteness or so. A certain object must have *all of its quantum dimensions collapsed* by the time it is transferred, or, better explained, *a certain object will be transferred if and only if it has all of its quantum dimensions collapsed*, so if in the merge position not all dimensions have collapsed because of the characteristics of the nearest head, the element must look for a procedural head that can license that dimension *prior to Transfer*. It is possible that Movement of [X] applies as a *last resort* to create a structural configuration where its dimensions can collapse while respecting Minimality, when the derivation has reached a point in which there is an intermediate element blocking the minimal configuration required for licensing / collapsing quantum dimensions in the *smallest unit* leading to convergence. This movement would be very much like Grohmann's (2003, 2004) *intra-clausal movement*, it goes upwards to the *immediately "dominant" informational domain*. The relevant syntactic object would not need to stop in each available position, but only on those that, once the relevant head is merged, can collapse the quantum dimension, as the derivation proceeds *in real time* and no object can "see" what will be merged later on. This is why we have *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 would imply maintaining a larger structure in the working memory, which is far from optimal. Movement applies as soon as it can (i.e., as soon as a new informational domain is created and we have 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 informational domain has the procedural features to make a certain dimension collapse, but not other, so *intra-domain* movement would be trivial, with no effect in the interface. If this is correct, then we will have found a way of making movement a principled operation, since there is no stipulation in its motivation (as there is if a feature like EPP is posited), only the dynamics of Merge (internal and external) in interaction with the interface requirements. The possibility is currently under research, since it is

possible that *quantum driven movement* applies in the C-I_1-syntax interface to reorganize elements of the RSS if needed to create appropriate configurations for collapse.

It is important to the analysis of *Transfer* to try to determine exactly the number and nature of the dimensions we are dealing with, since "full interpretability" means that all quantum dimensions have collapsed in a local relation with a functional head. We cannot do away with [Case], since it is important to C-I in the construction of the explicature and the interpretation of semantic functions (i.e., so-called "theta roles", see Krivochen, 2010c) The same happens, apparently, with dimensions like [definiteness] (or the procedural dimension one assigns D): a definite DP is fully interpretable by Relevance Theory's inferential component (i.e., C-I) and can undergo *referent assignment* (and is therefore a *phase*, transferred as soon as it can), an indefinite DP cannot in the usual cases, and it must wait until TP, AspP and ModP are merged and all the relevant dimensions have been licensed in the *phase level*. Therefore, this dimension seems relevant to determine transfer point, and cannot be dispensed with, along with T, Asp and Mod, of course. T gives a (caused or uncaused) generic event *reference*, Asp expresses the decision of the speaker as regards presenting the already delimited event as perfective or imperfective. Mod, in turn, represents the attitude of the speaker towards the proposition, in basic terms of realis / irrealis. In Bally's terms, the *dictum* (i.e., the *proposition*) would include AspP and its domain, to which a "truth value" can be assigned, and ModP would be the equivalent to the *modus*, which is out of the scope of truth evaluation in logical terms[21].

Now we have to justify the presence of the primitive dimensions [CAUSE], [EVENT] and [LOCATION] in the syntax, which seem to have been carried from earlier stages, if we follow the architecture we have proposed in Krivochen (2010b, d). In order to do this, we would need to reformulate the conservation principle, now that we have dispensed with features as they were traditionally conceived and any attempt of "valuation" process:

Revised formulation of the Conservation Principle:

a) *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.*

The aforementioned dimensions are the skeleton of the RSS, built in the pre-syntactic instance of C-I. These dimensions relate conceptual elements, and the whole RSS is not linguistic, as we hope to have demonstrated in

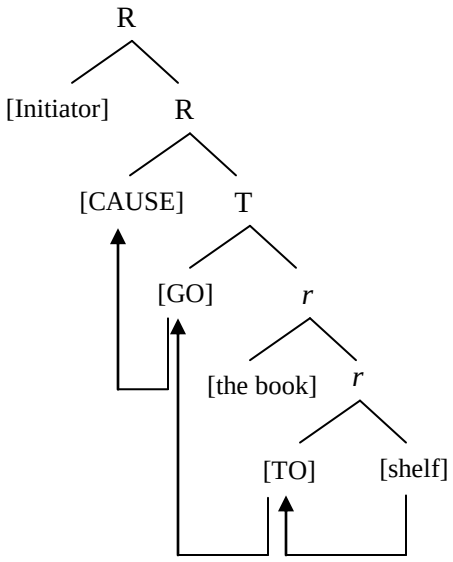
[21] In RT, the LF that the syntax generates as output is incapable of being assigned a truth value, since it is an incomplete propositional form (with referential expressions, vague elements, etc.). A fully-fledged explicature, on the contrary, can be truth-evaluated. That is the concept we are appealing to here, without commitment to one or other theory of truth.

Krivochen (2010d), but meaningful. RSSs are used to structure thought, and there is access to C-I without any participation of FL. For example, when we *see* a certain event taking place in the phenomenological world, we organize the participants and the relations in terms of a RSS which can be *unaccusative*, *unergative* or (di)*transitive* (see Mateu, 2000a, b), there is direct transference from the visual faculty to C-I. The dimensions, then, are *not linguistic*, and therefore cannot be read directly by FL but after an interface level, in which the instantiation of generic concepts into roots takes place. We have taken that interface to be DM's A List, which contains *roots* and *quantum dimensions*. Roots are severely *semantically underspecified* and *pre*-categorial, so they must be "assigned a category" (this terminology will be reviewed below) to generate an LF legitimate representation[22]. In Krivochen (2010b, d) we have posited the existence of a categorizer *c* comprising three dimensions, [Cause], [Event] and [Entity] (following a tendency that started in Marantz, 1997 and can be tracked up to Panagiotidis, 2010 and even more recent works). That categorizer, once merged with the relevant root, determined the category depending on the combination of the values of the dimensions. For example, an unergative verb was (l-syntactically) composed by a *root* merged with a *c* [+ Cause] [+ Event] [- Entity][23]. However, we can simplify this scenario with the framework outlined in this paper. We will take all roots to carry these dimensions as quantum dimensions. Roots, as such, in isolation, would still be *semantically underspecified* and thus "uninterpretable" (in the sense that they would make no contribution to the explicature) in C-I (Panagiotidis, 2009, 2010), but ***as linguistic entities they must have the inner potentiality of collapsing to a certain category***, if not, there would be nothing to justify the distinction between *roots* and *generic concepts* (or *conceptual addresses* and *concepts*, in Boeckx's terms). Remember that all we have to care about is *verbs* and *nouns*, since Ps are functional (that is, not associated with a lexical root, following the lines outlined in Krivochen 2010a) and Adj. and Adv. are both result of a conflation process (Mateu, 2000a, b), differing only in the type of elements they modify. How do we make these dimensions collapse? The answer is already predictable: by merging a functional head. Suppose we have a root with the quantum dimensions in the ψ-state. As such, it cannot be manipulated by FL, because it amounts to saying that it has no category, for practical purposes. We introduce that root in the working area and the functional category it merges with makes the quantum dimension collapse to one of the possible states, namely, *noun*, *derived nominal* (from caused and

[22] This is consistent with DM's "Categorization Assumption", which bans "bare" (i.e., uncategorized) roots in the derivation, but, as the reader will see, it is not the same, and the theoretical implications we draw are quite different. We are not invoking any principle or filter (which would be stipulative), only bare output conditions in the semantic component.

[23] Unergative verbs and transitive verbs share the categorizer values, but we have to take into account that unergative verbs are really transitive with a conflated nominal element and that superficial monotransitivity is really ditransitivity with a conflated argument (as in [shelve] or [saddle]). These transitive verbs are formed by adding a [Cause] relation to an unaccusative structure (thus yielding a [*v*P [VP [PP]]] structure), so that there would really be only two fundamental RSSs (and, therefore, two ways of conceptualizing the world): the *unergative* (caused, without spatial relation) and the *unaccusative* (uncaused, with spatial relation).

uncaused verbs), *caused verb*, *uncaused verb*. We have, then, a categorizer-less architecture, which is clearly simpler than a system that requires them, and we do not have to make big changes to the model outlined in previous papers, all we have to do is put the dimensions that *c* comprised within the root itself. However, the reader may remember that in earlier works we have strongly argued against categorizer-less systems, with the following argument: there is no way of avoiding the generation of bundles of features (*morphemes* in classic DM) that will collapse in the interfaces, terminal nodes to which no vocabulary item could correspond, taking as valid that Merge is unbounded and freely applied. That is, there is no (principled) way of avoiding overgeneration, which in turn results in a computational overload in the interface levels as the number of possibilities of combination is enormous even with a relatively small "lexical array". Even though we stick to that objection in a standard system (i.e., orthodox minimalism), our use of the *Conservation Principle* helps us have a non-stipulative way of restricting Merge and thus invalidate the objection. If the information has to be preserved *all-the-way-down* the derivational path (from the pre-syntactic instance of C-I to the external systems that receive the syntactic derivation as input) because of the *Conservation Principle*, then only that merger of a root with a functional category that results in the maintenance of information will be allowed by principle. Thus, if we start with a pragmatic global plan (in Bernárdez's terms) that "shapes" a RSS like the following (using Mateu's labels):



We would be conceptually generating a *location* V, namely, [shelve]. In order to preserve the information, we have to merge the *root* that instantiates the *complex generic concept* √SHEL- with an [eventive] node, a [cause] primitive and then Time, Asp and Mod to have a fully-fledged proposition including *dictum* and *modus*. In that

merger, the structural configuration for "category identification"[24] in LF is created, and the derivation converges. To make a generalization, a "verb" would arise from the following configuration:

i) [Time…α…√]

Being [α] an X number of non-intervenient nodes for categorization purposes, like {cause} or {event}, which can appear in nominal structures as well (for example, in derived nominals). Categorization, then, would depend on a local relation with a procedural head whose distribution is specified enough to determine an unambiguous interpretation from the structure that reaches the interface level. Our claim is that T is to V what D is to N, any root under the domain of D (in a minimal configuration) will be interpreted as a N.

This is currently under investigation, but could lead to a much simpler categorization system than current proposals (XSM, DM).

We may be accused in this point of abandoning *restrictivism* and arguing in favor of a *constructivist* theory that only generates convergent derivations (a "bad" version of *crash-proof*[25]). However, this is not so, because we are not proposing any specific syntactic filter or principle that restricts Merge in the working area (like s-/c-selection or subcategorization frames, for example), but invoking a much more basic principle of physical systems that assures the conservation of information throughout the derivation. *Syntax* remains the free unbounded mechanism of recursive combinatory that it has always been, but we have to bear in mind that, ***as part of the natural world, it is subjected to the same constraints and principles as any other physical system*** and ***it is not totally independent but restricted by interface conditions***.

Conclusion: a brief summary of Radical Minimalism

In this section we would like to list some basic tenets of Radical Minimalism (some of which are shared with the traditional Chomskyan view), to give it a more formal presentation and start outlining the model that we will, hopefully, expand in future papers. Our basic assumptions, and some conclusions we have drawn from them are:

[24] We use the term "category *identification*" as parallel to "label identification", as both are only relevant in the explicature level. In the syntax, therefore, there is no "category creation" in the sense that an a-categorial element is somehow *given* a category (for example, by merging it with a FC or a categorizer), as much as there is no "label creation", as argumented by Gallego (2010). The idea is that the "category" is interpreted from a configurational relation.

[25] Michael Putnam, p.c.

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*, SE, 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.*

5. From SRMT, we can follow the regularities that have been found regarding the structure of each faculty in the context of a massively modular mind (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).

Let us take as the last example EST phrase structure rules, which were *rewriting rules*. They were abandoned because of two reasons, only one of which was clear at that time: there were redundancies between the information they conveyed and the information present in the lexicon and because rewriting rules could tell us nothing specific about language, since the patterns they generate can be found in many domains of the natural world (Uriagereka, 1998). We can generate VPs as well as Fibonacci sequences, or describe fractal patterns. That was seen as a reason to abandon PSR, but we see it as a supporting argument for our hypothesis: regularities indicate that we are dealing with physical systems (as defined above) that are all part of the same physical universe (or "natural world", as is more common to say in Generative Grammar), which, optimally, should be regulated by a small set of general principles and allow a small set of universal operations (like *Merge*) to be performed.

Bibliography:

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

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

Bernárdez, E. (1982) *Introducción a la Lingüística del Texto*. Madrid, Epasa-Calpe.

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

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

(2010) *Defeating Lexicocentrism*. lingBuzz/001130

Bong-Kim, S (2010) *Is the EPP Feature Really Necessary? No!* lingBuzz/001165

Bosque, I. y Gutiérrez Rexach, J. (2008) *Fundamentos de sintaxis formal* Madrid, Akal.

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

Chomsky, N.: (1965) *Aspectos de la teoría de la sintaxis* Barcelona, Gedisa (1999).

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

(1982) *La nueva sintaxis. Teoría de la rección y el ligamiento.* Barcelona, Paidós (1988)

(1994) *Bare Phrase Structure*. In Webelhuth (ed.) (1995) *Government and Binding Theory and the Minimalist Program*. Oxford, Blackwell.

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

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

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

(2002) *Indagaciones minimalistas.* In Chomsky, Catalá, Piera, et. al. (2002). "El lenguaje y la mente humana". Barcelona, Ariel.

(2004) *Beyond Explanatory Adequacy*. En Belleti (ed.) "Structures and Beyond", Oxford, OUP.

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

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

(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.

Comrie, B. (1976) *Aspect*. Cambridge, CUP.

(1985) *Tense.* Cambridge, CUP.

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. Cambridge, MA. MIT Press.

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.

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". 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.

(1997b) *The Basic Elements of Argument Structure*. Ms. MIT.

(2002) *Conflation*. Ms. MIT.

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

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

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

Hinzen, W. (2009) *Hierarchy, Merge and Truth*. In Piatelli Palmarini et. al. "Of Minds and Language". Oxford, OUP.

Huang, C.-T. J. (1982) *Logical Relations in Chinese and the Theory of Grammar*, PhD thesis, MIT.

Hornstein, N. & P. Pietroski (2009) *Basic Operations: Minimal Syntax-Semantics.* In "Catalan Journal of Linguistics" n°8

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

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

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

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. Unpublished.

(2010b) *Algunas notas sobre fases*. UNLP. Unpublished.

(2010c) *θ-Theory Revisited.* UNLP. LingBuzz 001240.

(2010d) *Prolegómenos a una teoría de la mente*. UNLP. Unpublished.

(2011) *On the Syntax of Raising Verbs*. UNLP. Unpublished.

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

Leonetti, M. y Escandell, M. V. (2004) *Semántica conceptual, semántica procedimental*, en Actas del V congreso de Lingüística General, Madrid, Arco.

Marantz, A. (1995) *The Minimalist Program.* In Webelhuth (ed.), (1995).

(1997) *No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon*. En A. Dimitriadis, L. Siegel, C. Surek-Clark, and A. Williams, eds., *Proceedings of the 21st Penn Linguistics Colloquium*, UPenn Working Papers in Linguistics, Philadelphia.

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.

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

Panagiotidis, P. (2008) *Categorial Features of Functional Heads.* Universidad de Chipre.

(2009) *Four Questions on Categorization of Roots.* Ms. University of Cyprus.

(2010a) *Categorial Features and Categorizers*. Borrador del 31 de marzo de 2010. Universidad de Chipre.

(2010b) *Functional Heads, Agree and Labels*. Ms. University of Cyprus.

Philosophy of Quantum Mechanics MIT course Spring 2005. Available at http://ocw.mit.edu/courses/linguistics-and-philosophy/ Downloaded on 16/7/2008

Rizzi, L. (2004) *Locality and Left Periphery*. Ms.

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

Sperber, D. y Wilson, D. (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.

Uriagereka, J. (1998) *Pies y Cabeza*. Madrid, Visor.

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

Yus, F. (2010) *Relevance Theory*. In: The Oxford Handbook of Linguistic Analysis. Eds. B. Heine and H. Narrog. Oxford: Oxford University Press, 679-701.