Rebuilding MP on a Logical Ground

Alain Lecomte
Calligramme-LORIA

Abstract

We provide a logical system to express Minimalist Grammars, which aims at being "minimal" in the sense that it contains the least rules and the simplest lexical entries as possible. By limiting ourselves to the proofs in that system that satisfy one constraint on hypotheses management, we simulate minimalist derivations. This system is an elaboration on previous works by Lecomte and Retoré which were based on a new use of the Lambek calculus, where words were no longer associated with formulae, but with axioms and where a proof of a sentence was a proof of the theorem $\vdash c$ instead of a proof of a sequent $\Gamma \vdash c$. Such a calculus which suffered from limitations is here replaced by a version of partially non commutative linear logic, due to P. de Groote, and we show that in this system, when we limit ourselves to special proofs, we may mimick *move* (besides of course *merge*) for A-movement as well as head-movement. Moreover we show that the use of second-order types, allowed by the categorial grammar tradition, brings solutions for linguistic problems, including expletives and unbounded dependencies. We think that such a system is legitimated by its strong logical foundation, the fact that complexity results can be exported and also by philosophical reasons, which are linked to the way we may conceive logics: not only as a *formalisation of reasoning* but as a theory of general objects.

1 Introduction

In in his "draft" paper of 1998 [4], entitled *Minimalist Inquiries: the Framework*. Chomsky starts from the problem of what he calls *design specifications*.

> Let us invent a evolutionary fable, keeping it highly simplified. Imagine some primate with the human mental architecture and sensorimotor apparatus in place, but bo language organ. It had our modes of perceptual organization, our propositional attitudes (beliefs, desires, hopes, fears,...) insofar as these are not mediated by language, perhaps a "language of thought" in Jerry Fodor'sense, but no way to express its thoughts by means of linguistic expressions, so that they remain largely inaccessible to it, and of course to others. Suppose some event reorganizes the brain in such a way as, in effect, to insert FL. To be usable, the new organ has to meet "legibility conditions".

The design problem consists in defining an optimal device which meets these legibility conditions. Of course if that device also satisfies all other empirical conditions too: acquisition, processing, neurology, language change, that would be an "ideal" situation. Chomsky notes that this is obviously very unlikely... He therefore looks for the minimal devices to add in order to satisfy all these constraints. Among these necessary devices, Chomsky argues for the absolute necessity of at least two fundamental operations: *Merge* and *Attract* (see also [3]). These operations must satisfy the Inclusiveness condition:

In the course of a derivation, no elements are introduced by C_{HL}

That means that all the "elements" (in fact *features*) must be provided with the lexical entries. Operations of C_{HL} can only manipulate them (delete them, duplicate them, permutate them...). Chomsky claims that

*I am grateful to Christian Retoré for many helpful discussions. The seminal idea to change the traditional way of using systems like the Lambek calculus by using sequents with empty antecedents for typing linguistic expressions is his own. Of course all errors or imperfections in this paper are mine.

Merge is obligatory for any language-like system: it takes object already constructed and forms from them a new object. *Attract* seems, at first sight, motivated by an "apparent imperfection" of human language:

> commonly, phrases are interpreted in positions other than those where they are heard, though in analogous expressions these positions are occupied, and interpreted under natural conditions of locality.

Chomsky calls this fact *the dislocation property*. Dislocation of α yields a chain $< \alpha, \alpha >$. The raised element c-commands its trace in the original position[1]. But this "imperfection" can be related to another one in such a way that, together, they resolve in only one imperfection. This second "imperfection" concerns the existence of *uninterpretable features* of lexical items, that violate the interpretability condition:

Lexical items have no features other than those interpreted at the interface, properties of sound and meaning

An example of such a feature is *structural Case*. In order to "minimalize" the number of devices to introduce, we are lead to the assumption that dislocation is simply due to the need for deleting uninterpretabe features. If this is true, features are attracted by so called *attractors* and attractors are deleted when they have attracted their matching feature, something that Chomsky calls *Suicidal Greed*. Let us immediately notice that, *à la lettre*, that means that some feature *consumes* another one and after deletes, and that, therefore, there is some foundation to the idea which leads to a *resource sensitive logics* subjacent to that mechanism.
In such a reasoning, Chomsky empirically "finds" some operations which seem fundamental to him, and he tries to define other operations (like *Move*) by composition of these "primitive" operations. But of course he does not adress the question whether these operations are really primitive and if there could not be other operations, more primitive, that could generate them by composition, exactly like he does for *Move*. This appears to be a matter of logical analysis rather than of empirical study. Nevertheless, there is no reason to think that because it is logical analysis, it has no value... Simply because, like Chomsky himself claims it, the study of FL (the faculty of language) requires a "descriptive technology" which is not yet provided by any "direct" exploration of the way language works. At that stage, all instruments of thought can help, and logics among them.

The system here presented is an elaboration on previous works by Lecomte and Retoré [13]. We essentially try to give linguistic (and philosophic?) justifications to such a system and we propose some new solutions for dealing with semantics, and with particular phenomena like remnant movements, head movements, expletives and long distance effects.

2 A logical analysis of Merge

Let us therefore go back to the intuitions leading to *Merge* and *Attract*. The simplest way of characterizing *Merge* is to consider that an object which already exists (either from the lexicon or by previous construction) has some property which can only be satisfied by another object: that will put these two objects together, and then, when the property will be satisfied, we shall consider it as "inactive". The simplest way to express this kind of property is by means of a feature F. We say that an object O_1 has a feature F to satisfy. If O_2 can satisfy this feature, we can say that it contains the complementary feature F'. O_1 and O_2 then merge, that entails that F and F' are no longer active. Maybe some other feature of the new object becomes active, and the process can go on. Let us replace F by the notation /F, and F' by the notation F. Because the simplest items we can design are mere lists of features, let us suppose O_1 be ϕ/F, and O_2 be Fψ, where ϕ and ψ are sublists of features, we can have something like:

[1] *Merge* : $\phi/F \ F\psi \vdash \phi \ \psi$

[1] But Chomsky does not exclude cases where this relation can be violated, as in "independent XP-dislocation, in which the step-by-step locality and c-command relations for α_j are obliterated at LF", like:

`[written t`$_j$`for children]`$_i$`, [those books]`$_j$` couldn't possibly be t`$_i$`]`

At this point, we must ask: what composition law puts together the features of a same lexical entry, and then, of a same syntactic object? It is natural to assume that it is a kind of *product*, that we can denote by $\bullet$. What properties has this product? It is probably *associative* simply because, as we said above, we want to have the simplest structures as we can, that means concerning features in an object, a structure either of a (multi-)set or of a list, and not a tree-structure for instance, that would result from non-associativity. It is probably non-commutative, because we know from previous works on generative grammar that the formation of certain types of object must precede the formation of others. For instance, we remember the principles of the $\overline{\mathtt{X}}$-theory according to which, from the bottom, lexical heads are at first merged with their complements, in order to give a 1-bar object, and then and only then, this object can combine with a specifier in order to give a 2-bar object. Such an ordering requires the features be ordered in a syntactical object and therefore the product be non commutative. The question of neutral elements can be postponed, with its associate question of the inverses[2]. With this product, we rewrite [1] as:

[2] *Merge* : $\phi/F, F \bullet \psi \vdash \phi \bullet \psi$

If ψ is empty, we get: $\phi/F, F \vdash \phi$, where we recognize the usual cancellation scheme in the very elementary categorial grammars (Ajduckiewicz-Bar Hillel ones). In fact, this case is sufficient. Why? Now that we have the operators / and $\bullet$, we may recall some of the rules in the usual sequent presentation which manage them.

$$\frac{\Delta \vdash A \quad \Gamma, B, \Gamma' \vdash C}{\Gamma, B/A, \Delta, \Gamma' \vdash C}[L/] \qquad \frac{\Gamma, A, B, \Gamma' \vdash C}{\Gamma, A \bullet B, \Gamma' \vdash C}[L\bullet]$$

These two rules are *Introduction to the left* for the two operators. In Natural Deduction, they are presented as:

$$\frac{A/B \quad B}{A}[/E] \qquad \frac{A \bullet B \quad \begin{matrix}[A,B]^1\\ \vdots \\ C\end{matrix}}{C}[\bullet E]^1$$

or:

$$\frac{\Gamma \vdash A/B \quad \Delta \vdash B}{\Gamma, \Delta \vdash A}[/E] \qquad \frac{\Gamma \vdash A \bullet B \quad \Delta, A, B, \Delta' \vdash C}{\Delta, \Gamma, \Delta' \vdash C}[\bullet E]$$

and they are called *elimination rules.*
These rules have of course their duals, expressed as *Introduction to the right* in the sequent presentation and *introduction rules* in Natural Deduction:

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash B/A}[R/] \qquad \frac{\Gamma \vdash A \quad \Delta \vdash B}{\Gamma, \Delta \vdash A \bullet B}[R\bullet]$$

$$\frac{\begin{matrix}\Gamma, [A]^1\\ \vdots \\ B\end{matrix}}{B/A}[/I]^1 \qquad \frac{A, B}{A \bullet B}[\bullet I]$$

We must also assume the *identity rule*[3]

$$A \vdash A$$

[2]see [7] for a linguistic model which uses all the group operations

[3]In the usual "tree"-presentation of Natural Deduction, such a rule cannot be apparent, it is only apparent in the sequent calculus and in the sequent presentation of ND.

and therefore also, as a kind of dual principle, the *cut rule*:

$$\dfrac{\Gamma \vdash A \quad \Delta, A, \Delta' \vdash C}{\Delta, \Gamma, \Delta' \vdash C}[cut]$$

Let us take the sequent [2] to prove with these rules, say in the sequent calculus, we get:

$$\dfrac{\dfrac{F \vdash F \qquad \dfrac{\phi \vdash \phi \quad \psi \vdash \psi}{\phi, \psi \vdash \phi \bullet \psi}[R\bullet]}{\phi/F, F, \psi \vdash \phi \bullet \psi}[L/]}{\phi/F, F \bullet \psi \vdash \phi \bullet \psi}[L\bullet]$$

or, in the Natural Deduction style:

$$\dfrac{F \bullet \psi \qquad \dfrac{\dfrac{\phi/F \quad [F]^1}{\phi}[/E] \qquad [\psi]^1}{\phi \bullet \psi}[\bullet I]}{\phi \bullet \psi}[\bullet E]^1$$

It is worth to notice here that **two** hypotheses are *discharged* at the same time by the rule [$\bullet$E]: this will be fruitful in future application of this rule to the analogue of *Move*.
This shows that *Merge* has in fact more primitive operations: rules for $\bullet$ and the elimination rule (or introduction to the left, in the sequent presentation) for $/$.
Readers familiar with the Lambek calculus ([12]) recognize here this calculus. Of course, they are waiting for the introduction of the reverse $\backslash$. Why such an introduction is unavoidable? It is obviously because we have assumed $\bullet$ non commutative. As it can be very easily seen, *residuation laws* hold in this calculus. We have:

$$A \bullet B \vdash C \Leftrightarrow A \vdash C/B$$

For reasons of symmetry, we get the operator $\backslash$, in such a way that:

$$A \bullet B \vdash C \Leftrightarrow B \vdash A\backslash C$$

The $\backslash$ operator is not superfluous. With the $/$, we were able to attach a syntactic object to the right of another. If we follow the LCA convention of Kayne ([9]), that corresponds to the attachment of a complement. Because it is assumed in this framework that trees are binary, if the head selects a new syntactic object, it will be attached on the left, and therefore we shall have necessarily to use this $\backslash$. That simply duplicates the two rules for $/$, leading to analogous rules for $\backslash$.
According to Chomsky, there are two instances of *Merge*: set-Merge and pair-Merge. Set-Merge is "symmetric", in the sense that the object which is constructed from α and β is $\{\alpha,\beta\}$ where either label can project. Pair-Merge is *asymmetric*: it corresponds to the notion: "A adjoins to B", which is inherently different from "B adjoins to A". In the line of Chomsky's thought, we can ask whether it is possible to treat these two Merge in the same operations. In adjunction, "the adjoined element α leaves the category type unchanged". It is of course exactly what happens in our framework if some object α has type $\phi\bullet$ A/A and another one, β, type A$\bullet\psi$. By "merging" the two objects in the set-merge sense, we get an object which has still the feature (category) A, and therefore can be possibly selected by any object γ/A, exactly like it was the case for the previous object β.

3 Phonological interpretations

What happens to interpretable features, like phonological ones? Actually, we are lea d in our present reasoning to depart from the Chomskyan conception of phonological features, for a very simple reason.

When using our Natural Deduction rules, we are in the following situation. At each step, we build a new object: this object is a list of features, but if we want really have a **derivational** approach, or more precisely a really **markovian** derivation, like Chomsky himself claims to want[4], then we have to forget the tree-structure itself which is below the root of the new object, which contains the result of the computation on features. Therefore at each step we have to rearrange the phonological features of the "premisses" objects, in order to produce a new structure of the phonological features.
This leads to consider phonological features as **labels**, not in the chomskyan sense (for Chomsky, like for Stabler, labels are the entire lists of features), but in the type-logical sense. This enriches our rules, in the following way (cf. [15] for such a systematic labelling[5]).

$$\frac{\Delta \vdash \beta : A \quad \Gamma, \gamma : B, \Gamma' \vdash \delta : C}{\Gamma, \alpha : B/A, \Delta, \Gamma' \vdash \delta[\alpha\beta/\gamma]C}[L/] \qquad \frac{\Delta \vdash \beta : A \quad \Gamma, \gamma : B, \Gamma' \vdash \delta : C}{\Gamma, \Delta, \alpha : A\backslash B, \Gamma' \vdash \delta[\beta\alpha/\gamma]C}[L/]$$

$$\frac{\Gamma, \alpha : A, \beta : B, \Gamma' \vdash C}{\Gamma, \alpha\beta : A \bullet B, \Gamma' \vdash C}[L\bullet] \qquad \frac{\Gamma \vdash \alpha : A \quad \Delta \vdash \beta : B}{\Gamma, \Delta \vdash \alpha\beta : A \bullet B}[R\bullet]$$

Identity rules are labelled in the following way:

$$\alpha : A \vdash \alpha : A \; [axiom] \qquad \frac{\Gamma \vdash \alpha : A \quad \Delta, x : A, \Delta' \vdash \gamma : C}{\Delta, \Gamma, \Delta' \vdash \gamma[\alpha/x] : C}[cut]$$

Of course, we shall treat later on semantic features in the same way, thus adopting a position which seems to match precisely the view according to which:

> we understand L to be a device that generates expressions EXP, EXP=<PHON, SEM>, where PHON provides the "instructions" for sensorimotor systems and SEM for systems of thought.

In effect, to assert that amounts precisely to consider PHON and SEM features differently with regards to uninterpretable features. If the ideal picture was the case, that means if there were only interpretable features, simply motivated by the functioning of the sensorimotor system and the thought system, we should have only expressions consisting in this kind of pairs, but uninterpretable features bring a **third** component, in such a way that we get expressions EXP, EXP=<PHON,UNINT, SEM>, where UNINT means "uninterpretable". Merge and further Move steps are based on the consumption of uninterpretable features (either by selection – and then "inactivity", which is equivalent to suppression – or by "suicidal greed" in the case of Attract) but at each step also, interpretable features are preserved **and combined**. Let us notice that this is not contradictory with the inclusiveness condition.
Rules for $\bullet$ have this particular drawback that, when starting from the bottom (like the sequent calculus is usually used), the determination of α and β is not deterministic: if we have an object (lexical or syntactic) F$\bullet\phi$ endowed with a phonological feature α and we wish to use it in a derivation (that precisely means that it occurs in the left-hand side of a sequent, and therefore subject to the [L $\bullet$]-rule), we have nothing telling us how to split the string α. For instance, we have several ways of rewriting the deduction of [2] (one for each splitting of α into α_1 and α_2):

$$\dfrac{\dfrac{\alpha_1 : F \vdash \alpha_1 : F \qquad \dfrac{\beta\alpha_1 : \phi \vdash \beta\alpha_1 : \phi \quad \alpha_2 : \psi \vdash \alpha_2 : \psi}{\beta\alpha_1 : \phi, \alpha_2 : \psi \vdash \beta\alpha_1\alpha_2 : \phi \bullet \psi}[R\bullet]}{\beta : \phi/F, \alpha_1 : F, \alpha_2 : \psi \vdash \beta\alpha : \phi \bullet \psi}[L/]}{\beta : \phi/F, \alpha : F \bullet \psi \vdash \beta\alpha : \phi \bullet \psi}[L\bullet]$$

But in fact it does not matter: what only interests us is the endowment of the resulting object, and every splitting of α into (α_1,α_2) gives of course the same result, simply because in a Merge step, a product is

[4] Chomsky [p.24] requires the determination of the label of a category be "markovian", *requiring no information about the derivation*

[5] G. Morrill founds this labelling on the string semantics for such a system, in terms of 'groupoid' algebra.

decomposed (F•ψ) and then recomposed ($\phi \bullet \psi$), and in both operations α_1 and α_2 remain adjacent. We can therefore adopt a convention telling that the phonological feature is always attached to the "active" feature, that amounts here to adopt the splitting of α into (α,ϵ) where ϵ denotes the empty string.
It can also be interesting to show how things work in the Natural Deduction style. Here the labelled rules are:

$$\frac{\alpha : A/B \quad \beta : B}{\alpha\beta : A}[/E] \qquad \frac{\alpha : A \bullet B \qquad \begin{matrix}[x : A, y : B]^1 \\ \vdots \\ \gamma : C\end{matrix}}{\gamma[\alpha/xy]C}[\bullet E]^1$$

$$\frac{\begin{matrix}\Gamma, [x : A]^1 \\ \vdots \\ \beta x : B\end{matrix}}{\beta : B/A}[/I]^1 \qquad \frac{\alpha : A, \beta : B}{\alpha\beta : A \bullet B}[\bullet I]$$

In the Natural Deduction style, *hypotheses* are explicit under the form of formulae inside square brackets with an index, labelled with variables (x, y, ...). These variables range here over the set of strings representing phonological features. In the [•E] rule, the constant string α substitutes for the concatenation xy of the two string-variables. It is important here to notice that these string variables must be adjacent in order to make a *continuous* string. Let us rewrite the proof of [2] in this Natural Deduction format.

$$\frac{\beta : F \bullet \psi \qquad \dfrac{\dfrac{\alpha : \phi/F \quad [x : F]^1}{\alpha x : \phi}[/E] \qquad [y : \psi]^1}{\alpha xy : \phi \bullet \psi}[\bullet I]}{\alpha\beta : \phi \bullet \psi}[\bullet E]^1$$

In the third presentation, Natural Deduction in sequent format, left-hand sides of sequents are hypotheses and premises mixed and right-hand sides are conclusions, we systematically employ variables to label hypotheses and constants to label premises and conclusions depending on no hypotheses.

$$\frac{\Gamma \vdash \alpha : A/B \quad \Delta \vdash \beta : B}{\Gamma, \Delta \vdash \alpha\beta : A}[/E] \qquad \frac{\Gamma \vdash \alpha : A \bullet B \quad \Delta, x : A, y : B, \Delta' \vdash \gamma : C}{\Delta, \Gamma, \Delta' \vdash \gamma[\alpha/xy] : C}[\bullet E]$$

$$\frac{\Gamma, x : A \vdash \alpha x : B}{\Gamma \vdash \alpha : B/A}[R/] \qquad \frac{\Gamma \vdash \alpha : A \quad \Delta \vdash \beta : B}{\Gamma, \Delta \vdash \alpha\beta : A \bullet B}[R\bullet]$$

The same deduction (for [2]) has now the following appearance:

$$\frac{\vdash \beta : F \bullet \psi \qquad \dfrac{\dfrac{\vdash \alpha : \phi/F \quad x : F \vdash x : F}{x : F \vdash \alpha x : \phi}[/E] \qquad y : \psi \vdash y : \psi}{x : F, y : \psi \vdash \alpha xy : \phi \bullet \psi}[\bullet I]}{\vdash \alpha\beta : \phi \bullet \psi}[\bullet E]$$

Thus adopting Natural Deduction presentation, we can assume syntactical and lexical objets to be analogues of *theorems* and *axioms* in a formal system. Lexical objects are *proper axioms* (that means non-logical ones) labelled with a phonological feature (and further on a semantical one). Syntactic objects are *theorems*, that means hypotheses free deduction relations. Let us try a short example, using simply Merge.

Lexicon:

$$\begin{array}{lll} reads & ::= & \vdash reads : ((\overline{\mathtt{k}}\backslash\mathtt{vp})/\mathtt{d}) \\ a & ::= & \vdash a : ((\mathtt{d}\bullet\overline{\mathtt{k}})/\mathtt{n}) \\ book & ::= & \vdash book : \mathtt{n} \end{array}$$

Building a syntactic object: [3]

$$\dfrac{\dfrac{\vdash a : ((\mathtt{d}\bullet\overline{\mathtt{k}})/\mathtt{n}) \quad \vdash book : \mathtt{n}}{\vdash a\ book : \mathtt{d}\bullet\overline{\mathtt{k}}}[/E] \qquad \dfrac{\dfrac{\vdash reads : ((\overline{\mathtt{k}}\backslash\mathtt{vp})/\mathtt{d}) \quad x:\mathtt{d}\vdash x:\mathtt{d}}{x:\mathtt{d}\vdash reads\ x : (\overline{\mathtt{k}}\backslash\mathtt{vp})}[/E] \quad y:\overline{\mathtt{k}}\vdash y:\overline{\mathtt{k}}}{x:\mathtt{d}, y:\overline{\mathtt{k}} \vdash reads\ xy : (\overline{\mathtt{k}}\backslash\mathtt{vp})\bullet\overline{\mathtt{k}}}[\bullet I]}{\vdash reads\ a\ book : (\overline{\mathtt{k}}\backslash\mathtt{vp})\bullet\overline{\mathtt{k}}}[\bullet E]$$

The object so obtained, as expected, has the phonological feature *reads a book*, and uninterpretable features: $(\overline{\mathtt{k}}\backslash\mathtt{vp})$, $\overline{\mathtt{k}}$, in this order. It therefore contains a feature $\overline{\mathtt{k}}$ and the corresponding *attractor* $\overline{\mathtt{k}}\backslash$, and the categorial feature $\mathtt{vp}$.
There remains to explain how the attractor can attract its corresponding feature and then commit suicide on itself.

4 A logical analysis of Attract+Move

As currently observed,

> Attraction (hence movement) is driven by the need to delete an uninterpretable feature F; call it the attractor.

A good example is provided by our previous one, which gives an object which has an uninterpretable feature $\overline{\mathtt{k}}\backslash$ to delete. At this point, *Attract* must construct a new object K, free of this uninterpretable feature. For that:

> The attractor F in the label L of the target β locates the closest F' in its domain, attracting it to the MLI of F.

In our example, the closest F' is obviously $\overline{\mathtt{k}}$, which belongs to the same object. We can therefore assume this $\overline{\mathtt{k}}$ attracted by $\overline{\mathtt{k}}\backslash$ such that both are deleted.
Suppose we had got as a concluding theorem:

$$reads\ a\ book : \overline{\mathtt{k}}\bullet(\overline{\mathtt{k}}\backslash\mathtt{vp})$$

We could have had trivially the following steps: ([4])

$$\dfrac{\vdash reads\ a\ book : \overline{\mathtt{k}}\bullet(\overline{\mathtt{k}}\backslash\mathtt{vp}) \qquad \dfrac{x:\overline{\mathtt{k}}\vdash x:\overline{\mathtt{k}} \quad y:(\overline{\mathtt{k}}\backslash\mathtt{vp})\vdash : (\overline{\mathtt{k}}\backslash\mathtt{vp})}{x:\overline{\mathtt{k}}, y:(\overline{\mathtt{k}}\backslash\mathtt{vp})\vdash xy : \mathtt{vp}}[\backslash E]}{\vdash reads\ a\ book : \mathtt{vp}}[\bullet E]$$

There would have been no problem in this case because the factors of the product type would have been in the same order as the hypotheses to discharge, in conformity with the correct use of the [• E] rule. There would have been in fact no significative difference between *Merge* and *Abstract*. But it appears in most cases that phenomena are not so simple because the attracted feature is **not** at the right place where it can delete the attractor by means of our usual rules. In fact, if this very particular situation was general... *Attract* would not be *Attract* simply because the attracted feature would always been already attracted!
In the general case, the feature F' must be displaced in order to satisfy the requirement of the attractor.
This can only be done if we allow to *relax* the order of the hypotheses at some steps of our proofs. Such a *relaxation* is usually associated with the rule called *entropy rule* in most works on partially commutative linear logic ([6], [17], [1] ...). It enforces us to assume from now on that **two** products are acting: one, already

known and expressing order ($\bullet$), and the other one, to introduce, which ignores the order ($\otimes$). Both will be supposed associative. Mechanically, the assomption of a new, commutative, product $\otimes$ creates by the residuation technique a new "divisor", let us denote it $-\circ$ like in (commutative) Linear Logic ([8]). Moreover, we must recall that products are the reflects (internal to types) of structural operators used to compose the sequences of formulae on the left-hand side of the sequents (and also on the right-hand side when in a classical calculus). We must therefore assume from now on that two such structural operators can occur that we will denote:

$$\begin{array}{ll} , & for \otimes \\ ; & for \bullet \end{array}$$

We shall then assume the following structural rules:

$$\frac{\Gamma[(\Delta_1;\Delta_2)]\vdash A}{\Gamma[(\Delta_1,\Delta_2)]\vdash A}[entropy]$$

$$\frac{\Gamma[((\Delta_1,\Delta_2),\Delta_3)]\vdash A}{\Gamma[(\Delta_1,(\Delta_2,\Delta_3)))]\vdash A}[\otimes ass1] \qquad \frac{\Gamma[(\Delta_1,(\Delta_2,\Delta_3))]\vdash A}{\Gamma[((\Delta_1,\Delta_2),\Delta_3)]\vdash A}[\otimes ass2]$$

$$\frac{\Gamma[((\Delta_1;\Delta_2);\Delta_3)]\vdash A}{\Gamma[(\Delta_1;(\Delta_2;\Delta_3)))]\vdash A}[\bullet ass1] \qquad \frac{\Gamma[(\Delta_1;(\Delta_2;\Delta_3))]\vdash A}{\Gamma[((\Delta_1;\Delta_2);\Delta_3)]\vdash A}[\bullet ass2]$$

$$\frac{\Gamma[(\Delta_1,\Delta_2)]\vdash A}{\Gamma[(\Delta_2,\Delta_1)]\vdash A}[comm]$$

together with the following "logical" rules:

$$\frac{\Gamma\vdash\alpha:A/B \quad \Delta\vdash\beta:B}{\Gamma;\Delta\vdash\alpha\beta:A}[/E] \qquad \frac{\Gamma\vdash\alpha:A\bullet B \quad \Delta;x:A;y:B;\Delta'\vdash\gamma:C}{\Delta;\Gamma;\Delta'\vdash\gamma[\alpha/xy]:C}[\bullet E]$$

$$\frac{\Gamma;x:A\vdash\alpha x:B}{\Gamma\vdash\alpha:B/A}[/I] \qquad \frac{\Gamma\vdash\alpha:A \quad \Delta\vdash\beta:B}{\Gamma;\Delta\vdash\alpha\beta:A\bullet B}[\bullet I]$$

$$\frac{\Gamma\vdash\alpha:B-\circ A \quad \Delta\vdash\beta:B}{\Gamma,\Delta\vdash\{\alpha,\beta\}:A}[-\circ E] \qquad \frac{\Gamma\vdash\alpha:A\otimes B \quad \Delta,x:A,y:B\vdash\gamma:C}{\Gamma,\Delta\vdash\gamma[\alpha/\{x,y\}]:C}[\otimes E]$$

$$\frac{\Gamma,x:A\vdash\{\alpha,x\}:B}{\Gamma\vdash\alpha:B-\circ A}[-\circ I] \qquad \frac{\Gamma\vdash\alpha:A \quad \Delta\vdash\beta:B}{\Gamma,\Delta\vdash\{\alpha,\beta\}:A\otimes B}[\otimes I]$$

We have endowed these rules with phonological labels which respect commutativity and non commutativity. $\{\alpha,\beta\}$ means the "union" of the phonological features α and β without any order between them. [$-\circ$ E] and [$\otimes$ I] make such unions, [$-\circ$ I] makes a "substraction". The substitution proposed in [$\otimes$ E] seems at first sight mysterious: what is the substitution of α to the unordered set $\{x, y\}$? We shall assume here, as a convention on labelling, that it is the **simultaneous** substitution of α for both x and y, *no matter the order between x and y is.*
With a new lexicon, we can obtain new proofs for syntactic objects, like the following one.

Lexicon:

$$\begin{array}{lll} reads & ::= & \vdash reads : ((\overline{\mathtt{k}}\backslash\mathtt{vp})/\mathtt{d}) \\ a & ::= & \vdash a : ((\mathtt{d}\otimes\overline{\mathtt{k}})/\mathtt{n}) \\ book & ::= & \vdash book : \mathtt{n} \end{array}$$

Building a new syntactic object: [5]

$$\dfrac{\dfrac{\vdash a : ((\mathtt{d}\otimes\overline{\mathtt{k}})/\mathtt{n}) \quad \vdash \mathit{book} : \mathtt{n}}{\vdash a\ \mathit{book} : \mathtt{d}\otimes\overline{\mathtt{k}}}[/E] \quad \dfrac{\dfrac{y:\overline{\mathtt{k}}\vdash y:\overline{\mathtt{k}} \quad \dfrac{\vdash \mathit{reads} : ((\overline{\mathtt{k}}\backslash \mathtt{vp})/\mathtt{d}) \quad x:\mathtt{d}\vdash x:\mathtt{d}}{x:\mathtt{d}\vdash \mathit{reads}\ x : (\overline{\mathtt{k}}\backslash\mathtt{vp})}[/E]}{y:\overline{\mathtt{k}}; x:\mathtt{d}\vdash y\ \mathit{reads}\ x:\mathtt{vp}}[\backslash E]}{y:\overline{\mathtt{k}}, x:\mathtt{d}\vdash y\ \mathit{reads}\ x:\mathtt{vp}}[\mathit{entropy}]}{\vdash a\ \mathit{book}\ \mathit{reads}\ a\ \mathit{book} : \mathtt{vp}}[\otimes E]$$

What is striking here is the fact that we get an order in the label even if we have got through entropy. It is so because when using [/ E] and [\ E], we necessarily order the labels, and this order (inside the label) is never destroyed, even when using the entropy rule: at this moment, it is only the order on hypotheses which is relaxed and this relaxation allows only to discharge them by means of a $\otimes$-type, thus resulting, after our convention, in simultaneous substitution of the string labels x and y inside the string already built in the consequent.
If things are so, we are lead to the conclusion that order is created in that system by the connectives / and \, and that the product is mainly used for discharging hypotheses and is therefore required to be commutative. We can even ask whether the $\bullet$-product ever occurs in UNINT. We can perhaps completely dispense with our previous assomption, expressed in [2], according to which a lexical or syntactic object consists in a $\bullet$-product of features, something we denoted by F$\bullet\phi$... What we viewed in an intuitive, non formal setting, as a product of "attractors" $F_1\backslash, F_2\backslash$, ..., $F_n\backslash, /F_{n+1}$ including a categorial feature ϕ can obviously be expressed formally without any $\bullet$-product, simply writing: $((F_1\backslash(F_2\backslash...(F_n\backslash\phi)))/F_{n+1})$.
These observations lead us now to a real *simplification* of the system, because if it is so, we can get rid of the rules for $\bullet$. The deductions [3,4] will no longer belong to our set of proofs, thus eliminating an embarrassing fact of potential overgeneration. That means that, from now on, a proper subset of proofs will be selected inside our deductive system, those proofs will be said *admissible* relatively to some linguistic criteria. We set as an assumption:

Proposition 1 (**First criterion on admissible proofs**) *Admissible proofs contain only three kinds of steps:*

- *implication steps (elimination rules for / and \)*
- *tensor steps (elimination rule for $\otimes$)*
- *entropy steps (entropy rule)*

From this we conclude now that the operation *Merge* is much simpler than what we had expected in the first section: we can eliminate the [$\bullet$ I] and [$\bullet$ E] steps, and then, *Merge* consists in a simple *implication*-step[6].
Attract simply consists in applying the entropy rule to the set of hypotheses which appear on the left-hand side of our ND-sequents, and then possibly applying the commutativity rule in order to make the attracted feature close to its attractor, but this is in fact useless simply because [$\otimes$ E] can act on non adjacent features. Therefore *Attract* is limited to the operation *locate the closest F'*, something to which we will return in a moment. After *Attract*, F' must determine the phrase α, *a candidate for pied-piping* like Chomsky says, in order to finally make α merge with a category K. In our presentation, α is the phrase associated with the $\otimes$-product type which contains F' as a conjunct. Informally, it must merge with the syntactic object having F as an attractor, that means of the form F$\backslash\phi$ (if we still want that all moves are left- and up- ward).

[6] We could suspect the entropy rule be responsible of a loss of resource sensitivity: of course it is not the case with these restrictions because we cannot obtain in this subset any proof of the kind $A/B \vdash B\multimap A$, but even if we stand in the whole system, where we can get a proof of such a sequent, this is not a serious problem, firstly because we cannot get the reverse deduction relation and secondly because such a result *does not mean* that every object of type A/B and therefore searching a B on its right in order to become an A, is also an object insensitive to the direction where it looks for B, it only means that if an object is of type A/B, we of course know that it misses a B in order to become an A. Such a provable sequent is therefore only a weakening rule, it cannot be used for proving for instance *(B,A/B⊢A). In other terms, $B\multimap A$ is *underdetermined* with regards to A/B or B\A.

Formally, at the step where α, of type F$\otimes\psi$, merge with K of type F$\backslash\phi$, in fact a merge step has already been done: it was between K and a *hypothesis* x, of type F. The only thing which happens therefore after *Attract* is *discharging* the hypothesis x, by means of α. To say that F' must belong to the so-called domain D(F) of F is to say that, unformally speaking, the product object to which F' belongs has already merged during the construction of the object containing the attractor F. In our setting, this merge step has consisted in merging another hypothesis y at some point of the derivation, a hypothesis which is now discharged by means of α and [$\otimes$ E] simultaneously with x. Therefore *Move* is the complete application of [$\otimes$ E] after locating the correct attracted feature.
Let us return now to *Attract*. According to the *Greed* principle, an attractor wants to be satisfied *as soon as possible*! This is the same idea which is expressed in the *Shortest Move* condition and in the *Minimal Link Condition*. That simply translates in our setting into the following new admissibility condition:

Proposition 2 *(***Second criterion on admissible proofs***) Hypotheses must be discharged according to an order First In, First Out.*

We can take the metaphor of an *agenda* for the left hand side of a sequent: hypotheses are put on this agenda in their order of introduction (we can give them a priority order when they are introduced) and this indicates emergency tasks to do. As soon this agenda already containing a hypothesis x of type A includes some complementary hypothesis y of type B, that means a new hypothesis associated with a type B such that A$\otimes$B is the type of a lexical object or of a syntactic object that can be built, [$\otimes$ E] must apply, discharging simultaneously x and y.
It is interesting to note that, looking at the derivation tree, if x is introduced before y, this translates into the fact that x is *lower* than y in this tree. The alternative way of doing, if we want to avoid numbering hypotheses with priority order, is therefore to refer to this relation of *heighth* in the derivation tree. This has nevertheless the drawback of enforcing us to refer to the entire derivation and therefore abandoning the "markovian" principle, something that we wanted to avoid if possible.
Putting aside these "imperfections", we get a very simple model... perhaps the simplest we can imagine for doing so intricated tasks.

5 Rules and lexicon for a minimalist deductive system

To sum up, we give the small set of rules and the structure of lexical entries which seem to be sufficient for expressing minimalist principles. Let us call $\mathcal{MDS}$ (for Minimalist Deductive System) the fragment of pCILL (*partially Commutative Linear Logic*) we thus obtain.

Rules

$$\frac{\Gamma[(\Delta_1;\Delta_2)] \vdash A}{\Gamma[(\Delta_1,\Delta_2)] \vdash A}[entropy]$$

$$\frac{\Gamma \vdash \alpha : A/B \quad \Delta \vdash \beta : B}{\Gamma;\Delta \vdash \alpha\beta : A}[/E] \qquad\qquad \frac{\Delta \vdash \beta : B \quad \Gamma \vdash \alpha : B\backslash A}{\Delta;\Gamma \vdash \alpha\beta : A}[\backslash E]$$

$$\frac{\Gamma \vdash \alpha : A \otimes B \quad \Delta, x : A, y : B \vdash \gamma : C}{\Gamma, \Delta \vdash \gamma[\alpha/\{x,y\}] : C}[\otimes E]$$

to which we must add as criterion for admissible proofs:

Hypotheses must be discharged according to the order First In, First Out.

Lexicon:

A lexical entry consists in an axiom $\vdash w : \mathcal{T}$ where $\mathcal{T}$ is a type:

$$((F_2\backslash(F_3\backslash...(F_n\backslash(G_1 \otimes G_2 \otimes ... \otimes G_m \otimes A))))/F_1)$$

where:

- m and n can be any number greater than or equal to 0,
- F_1, ..., F_n are *attractors*,
- G_1, ..., G_m are features,
- A is the resulting category type

Let us notice that if n=0, we simply get a $\otimes$-product type, if n=1, we get a type which attracts a phrase on its right (thus giving a *complement*), if n>1, the phrases are all attracted on the left, except for the first attractor (thus giving *specifiors*).

6 Semantic interpretations

6.1 Labelling types

As indicated above, we consider objects being associated with tuples <PHON, UNINT, SEM>. The most natural solution for including semantical features is to treat them as another kind of labels on which we shall return later on.
Borrowing from works inside the categorial or the type-theoretical framework ([14], [16], [15]), we simply assume that semantical representations are *given by proofs*[7]. We can now see that the object that we build step by step in such a deductive system and was assumed to be of no use on the **syntactic** side (except for keeping track of the order of hypotheses) is in fact *the object which gives us the semantics of the constructed object*. Actually, it will not give directly such a semantics, this semantics will be extracted from the proof-tree in a straightforward way.
Semantics is here assumed to be **compositional**. It is the reason why we aim at treating it by means of *application* steps and *abstraction* steps, like in ordinary lambda-calculus. Compositionality seems in effect to be the least requirement we can have, in order to have a correct representation of phenomena like *scoping*, *coreferentiality*, *predicate argument structure*, *quantificational structure* etc. all things being absolutely necessary to feed up the systems of thought.
Merge is obviously associated with *application*. Let us imagine for instance an object which would be built up by means of sole *Merge*.

Lexicon:

$$\begin{array}{lll} \textit{loves} & ::= & \vdash \textit{loves} : (\mathtt{vp}/\mathtt{d}) : (\mathbf{l}) \\ \textit{mary} & ::= & \vdash \textit{mary} : \mathtt{d} : (\mathbf{m}) \end{array}$$

Building an object:

$$\frac{\vdash \textit{loves} : (\mathtt{vp}/\mathtt{d}) : (\mathbf{l}) \quad \vdash \textit{mary} : \mathtt{d} : (\mathbf{m})}{\vdash \textit{loves mary} : \mathtt{vp} : (\mathbf{l}(\mathbf{m}))} [/E]$$

We can simply say that the term we get encodes the proof (borrowing from the famous Curry-Howard homomorphism between intuitionistic proofs and lambda terms ([5])).

Things are more difficult with *Move*. In fact, in a *Move* step, an object which has already merged (under the instance of a variable x) re-merges a second time. Using application associated with *Merge*, that would

[7] This belongs in fact to a long tradition in Intuitionistic Logic, which dates back to earlier works by Brouwer, Heyting... and is also known as the Brouwer-Heyting-Kolomgorov assumption.

entail the application of a semantics S_1 to a semantics S_2, which already comes from the application of a semantics S to S_1!

We can solve this apparent paradox by associating with a $\otimes$-product F$\otimes$G were F merges at a *higher* position than G, a vector (ϕ, ξ), where ϕ is associated with F and ξ with G. ξ is a semantic variable[8]. But what is ϕ? In fact as noted by Stabler ([19]), a raised constituent has also a raised semantic type. But in its second occurrence, "raised" means that it is *lifted* in the montagovian sense. For instance, instead of representing the item *Mary* by the constant `mary`, we represent it by the higher order formula $\lambda P.P(mary)$. The semantics S_1 in this case is splitted, in such a way that we have during the first merge step: application of some semantics S to ξ, giving S_2, and in a second merge, application of the lifted semantics in S_1, say here for instance: $\lambda P.P(mary)$ to S_2. But this requires S_2 be a predicate, or a lambda term of the form $\lambda x.Q(x)$, where x is an individual variable, and in fact we have generally got at this place a saturated predicate $Q(\xi)$.

The way to solve this problem of conflicting types is to perform an *abstraction* step just before the higher order formula applies, involving an abstraction precisely on the variable ξ wich now occurs in the body of the formula.

Moreover, "syntactic" proofs are transformed into "semantic" ones for the following reasons: the semantic world is insensitive to the direction in which arguments are found and therefore the distinction between / and \ collapses. Only $\multimap$ is used. For the time being, we don't know whether any product (with its correlatives *pairing* and *projections*) has a role to play on this semantic side and therefore we shall dispense with them. Finally and more importantly, a referential semantics works only with types denoting *referential entities*($\mathbf{e}$) and *truth values* ($\mathbf{t}$) and all functional entities we can build from them by means of the only connective $\multimap$[9]. The extracted semantic proof therefore partly results from a translation of syntactic types into semantic ones, a translation that we can see as a homomorphism $\mathcal{H}$:

- $\mathcal{H}(\mathtt{d}) = \mathbf{e}$
- $\mathcal{H}(\mathtt{vp}) \in \{\mathbf{t}, (\mathbf{e}\multimap\mathbf{t}), (\mathbf{e}\multimap(\mathbf{e}\multimap\mathbf{t}))...\}$
- $\mathcal{H}(\overline{\mathtt{f}}) \in \{\mathbf{e}, ((\mathbf{e}\multimap\mathbf{X})\multimap\mathbf{X})\}$, **where** $X \in \{\mathbf{t}, (\mathbf{e}\multimap\mathbf{t}), (\mathbf{e}\multimap(\mathbf{e}\multimap\mathbf{t}))...\}$
- $\mathcal{H}(K/K') = \mathcal{H}(K'\backslash K) = (\mathcal{H}(K')\multimap\mathcal{H}(K))$

We may see therefore that there is no room for any [$\otimes$ E] or [$\bullet$ E] in image proofs, but there still remains hypotheses labelled with variables, hypotheses that can only be discharged by [$\multimap$ I] steps. Entropy steps are useless and therefore suppressed. [/ E] and [\ E] steps are replaced by [$\multimap$ E] steps. Because hypotheses must be discharged, they are discharged just before [$\multimap$ E] steps consisting in applying the semantics of the *coindexed* higher order feature, or if there are other coindexed hypotheses, just before [$\multimap$ E] steps consisting in cancelling variables associated with them.

We moreover assume that all variables labelling hypotheses in syntactic proofs are substituted by the corresponding semantic components of the product types used to discharge them in the syntactic proof. Occurrences of product types are therefore suppressed. Let us see for instance what happens with the very short example given by *loves Mary*:

Lexicon:

$$\begin{array}{lll} loves & ::= & \vdash loves : ((\overline{\mathtt{k}}\backslash\mathtt{vp})/\mathtt{d}) : (\lambda x.\mathbf{l}(x)) \\ mary & ::= & \vdash mary : \overline{\mathtt{k}} \otimes \mathtt{d} : (\lambda P.P(\mathbf{m}), \xi) \end{array}$$

Building a syntactic object:

[8] i.e. a variable which occurs in the semantic component

[9] or perhaps $\rightarrow$ if we find that the semantic world is also insensitive to the amount of resources, but in this case, the problem will be solved siply by adding the unary connective **!** of Linear Logic, according to the well known decomposition: $A \rightarrow B \equiv !A\multimap B$.

$$\dfrac{\vdash \mathit{mary} : \mathtt{d} \otimes \overline{\mathtt{k}} \qquad \dfrac{\dfrac{y : \overline{\mathtt{k}} \vdash y : \overline{\mathtt{k}} \qquad \dfrac{\vdash \mathit{loves} : ((\overline{\mathtt{k}}\backslash \mathtt{vp})/\mathtt{d}) \quad x : \mathtt{d} \vdash x : \mathtt{d}}{x : \mathtt{d} \vdash \mathit{loves}\ x : (\overline{\mathtt{k}}\backslash \mathtt{vp})}[/E]}{y : \overline{\mathtt{k}}; x : \mathtt{d} \vdash y\ \mathit{loves}\ x : \mathtt{vp}}[\backslash E]}{y : \overline{\mathtt{k}}, x : \mathtt{d} \vdash y\ \mathit{loves}\ x : \mathtt{vp}}[\mathit{entropy}]}{\vdash \mathit{mary\ loves\ mary} : \mathtt{vp}}[\otimes E]$$

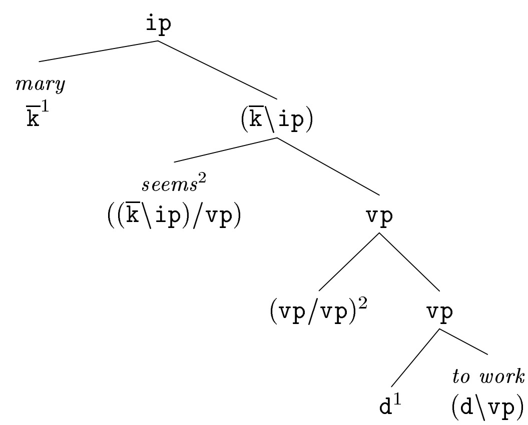
Building the corresponding semantic object:

$$\dfrac{\vdash \lambda P.P(\mathbf{m}) : ((\mathbf{e}\multimap\mathbf{t})\multimap\mathbf{t}) \qquad \dfrac{\dfrac{\vdash \lambda x.\mathbf{l}(x) : (\mathbf{e}\multimap\mathbf{t}) \quad \mathbf{x} : \mathbf{e} \vdash \mathbf{x} : \mathbf{e}}{x : \mathbf{e} \vdash \mathbf{l}(\mathbf{x}) : \mathbf{t}}[\multimap E]}{\vdash \lambda x.\mathbf{l}(x) : (\mathbf{e}\multimap\mathbf{t})}[\multimap I]}{\vdash \mathbf{l}(\mathbf{m}) : \mathtt{vp}}[\multimap E]$$

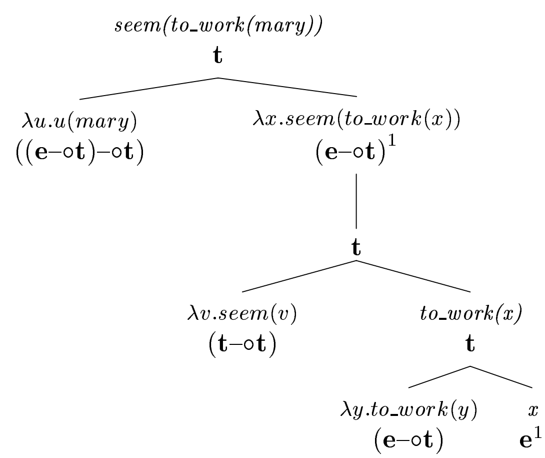
The following example concerns subject-raising. we assume:

mary	::=	$\vdash \mathit{mary} : \overline{\mathtt{k}} \otimes \mathtt{d} : (\lambda u.u(\mathit{mary}), x)$
$\mathtt{seems}$	::=	$\vdash \mathit{seems} : ((\overline{\mathtt{k}}\backslash\mathtt{ip})/\mathtt{vp}) \otimes (\mathtt{vp}/\mathtt{vp}) : (\emptyset, \lambda v.\mathit{seem}(v))$
$\mathtt{to\_work}$	::=	$\vdash \mathit{to\ work} : (\mathtt{d}\backslash\mathtt{vp}) : \lambda y.\mathit{to\_work}(y)$

We can represent the derivation of *Mary seems to work* by the following tree (thus using the natural deduction presentation, put upside down):



where indices refer to hypotheses which are discharged altogether. Then, eliminating semantically empty nodes and transforming the tree along the lines stated above, we obtain:



6.2 Note on the interpretative components

It is worth to notice here that in order to produce phonological and logical forms, we have not yet used the distinction between weak and strong features, as done in previous works by Stabler [18]. In most cases, it suffices in fact to label not the entire types themselves, but their factors. It is what was done in the previous subsection by using vectors associated with product types, but the same could have been done for phonological interpretations.
Unfortunately, this cannot be generalized: in a SVO language, the DP-subject is overtly raised and the DP-object is only covertly raised, and we cannot have separate entries for DP-subjects and DP-objects. This legitimates a kind of distinction concerning only uninterpreted features like `case`. In the phonological component, the *attractor* associated with such a feature must indicate whether it requires a plain phonology or an empty one. We shall write $\overline{\mathtt{k}}^{*}$ for the first case, and the lexical entries will be labelled not by vectors, but by multi-sets, the choice of the order being made only at the discharging step. In case of phonology, nodes which will not receive any content will be labelled by ϵ, in case of semantics, they will be simply deleted.
This requires a more complete formulation of the [$\otimes$ E] law, which will be stated further.

7 Unsolved problems and perspectives

7.1 Remnant Movement

In some new works in the Chomskyan paradigm (cf. Chomsky himself in the quoted draft, but also Koopman and Szabolcsi ([10]), Stabler([19])), Head Movement seems to be useless. For instance, Stabler argues in favour of grammars without head movement and without covert movement, thus promoting *remnant movements*, where a "remnant" is:

> a constituent from which material has been extracted

and Stabler adds:

> Moving a constituent from which material has already been extracted means that traces of earlier movements may be carried to positions where they are no longer c-commanded by their antecedents, something which was banned in earlier theories.

Remnant movement is mainly motivated by examples from Hungarian. "Developing an observation of Kenesei, Koopman and Szabolcsi observe the folloowing pattern in negated or focused sentences of Hungarian, schematized on the right where "M" is used to represent the special category of verbal modifiers like *haza-*":

(1) Nem fogok akarni kezdeni haza-menni [V1 V2 V3 M V4]
(not will want begin home-go)

(2) Nem fogok akarni haza-menni kezdeni [V1 V2 M V4 V3]
(not will want home-go begin)

(3) Nem fogok haza-menni kezdeni akarni [V1 M V4 V3 V2]
(not will home-go begin want)

Stabler shows that it is possible to have an account of such phenomena by giving lexical entries features which trigger the "rolling-up" of verb phrases. The features he gives can be translated in our framework by:

Lexicon:

$$\begin{array}{lll}
V1 & ::= & \vdash V1 : (\mathtt{v}/\mathtt{v}) \\
V2 & ::= & \vdash V2 : (\mathtt{v}/\mathtt{v}) \\
V2 & ::= & \vdash V2 : ((\overline{\mathtt{v}}\backslash\mathtt{v})/\mathtt{v}) \\
V3 & ::= & \vdash V3 : (\mathtt{v}/\mathtt{v}) \\
V3 & ::= & \vdash V3 : ((\overline{\mathtt{v}}\backslash\mathtt{v})/\mathtt{v}) \\
V3 & ::= & \vdash V3 : ((\overline{\mathtt{v}}\backslash(\mathtt{v}\otimes\overline{\mathtt{v}}))/\mathtt{v}) \\
V4 & ::= & \vdash V4 : ((\overline{\mathtt{m}}\backslash\mathtt{v})/\mathtt{m}) \\
V4 & ::= & \vdash V4 : ((\overline{\mathtt{m}}\backslash(\mathtt{v}\otimes\overline{\mathtt{v}}))/\mathtt{m}) \\
M & ::= & \vdash M : \mathtt{m}\otimes\overline{\mathtt{m}}
\end{array}$$

A proof giving the correct order [V1 V2 V3 M V4] is the following (in tree-format):

$$\dfrac{M : \mathtt{m}\otimes\overline{\mathtt{m}} \quad \dfrac{V1 : (\mathtt{v}/\mathtt{v}) \quad \dfrac{V2 : (\mathtt{v}/\mathtt{v}) \quad \dfrac{V3 : (\mathtt{v}/\mathtt{v}) \quad \dfrac{y : \overline{\mathtt{m}} \quad \dfrac{V4 : ((\overline{\mathtt{m}}\backslash(\mathtt{v}\otimes\overline{\mathtt{v}}))/\mathtt{m}) \quad x : \mathtt{m}}{V4x : (\overline{\mathtt{m}}\backslash(\mathtt{v}\otimes\overline{\mathtt{v}}))}}{yV4x : \mathtt{v}}}{V3yV4x : \mathtt{v}}}{V2V3yV4x : \mathtt{v}}}{V1V2V3yV4x : \mathtt{v}}}{V1V2V3MV4 : \mathtt{v}}$$

A proof giving the correct order [V1 M V4 V3 V2] is the following:

$$\dfrac{\dfrac{\dfrac{M : \mathtt{m}\otimes\overline{\mathtt{m}} \quad \dfrac{y : \mathtt{m} \quad \dfrac{V4 : ((\mathtt{m}\backslash(\mathtt{v}\otimes\overline{\mathtt{v}}))/\overline{\mathtt{m}}) \quad x : \mathtt{m}}{V4x : (\mathtt{m}\backslash(\mathtt{v}\otimes\overline{\mathtt{v}}))}}{yV4x : \mathtt{v}\otimes\overline{\mathtt{v}}}}{MV4 : \mathtt{v}\otimes\overline{\mathtt{v}}} \quad \dfrac{x' : \overline{\mathtt{v}} \quad \dfrac{V3 : ((\overline{\mathtt{v}}\backslash(\mathtt{v}\otimes\overline{\mathtt{v}}))/\mathtt{v}) \quad y' : \mathtt{v}}{V3y' : (\overline{\mathtt{v}}\backslash(\mathtt{v}\otimes\overline{\mathtt{v}}))}}{x'V3y' : \mathtt{v}\otimes\overline{\mathtt{v}}}}{MV4V3 : \mathtt{v}\otimes\overline{\mathtt{v}}} \quad \dfrac{V1 : (\mathtt{v}/\mathtt{v}) \quad \dfrac{x'' : \overline{\mathtt{v}} \quad \dfrac{V2 : ((\overline{\mathtt{v}}\backslash\mathtt{v})/\mathtt{v}) \quad y'' : \mathtt{v}}{V2y'' : (\overline{\mathtt{v}}\backslash\mathtt{v})}}{x''V2y'' : \mathtt{v}}}{V1x''V2y'' : \mathtt{v}}}{V1MV4V3V2 : \mathtt{v}}$$

If we analyze such a proof-system, we perceive that:

that $V4$ be $((\overline{\mathtt{m}}\backslash(\mathtt{v}\otimes\overline{\mathtt{v}}))/\mathtt{m})$ and M be $\mathtt{m}\otimes\overline{\mathtt{m}}$ entails the inversion of M and $V4$, resulting in $MV4 : \mathtt{v}\otimes\overline{\mathtt{v}}$,

that $V3$ be $((\overline{\mathtt{v}}\backslash(\mathtt{v}\otimes\overline{\mathtt{v}}))/\mathtt{v})$ and $MV4$ be $\mathtt{v}\otimes\overline{\mathtt{v}}$ entails the inversion of $MV4$ and $V3$, resulting in $MV4V3$ which is still $MV4 : \mathtt{v}\otimes\overline{\mathtt{v}}$, (let us note that therefore, at this stage, the same processus could ..undefinitely be continued)

that $V2$ be $((\overline{\mathtt{v}}\backslash\mathtt{v})/\mathtt{v})$ and $MV4V3$ be $\mathtt{v}\otimes\overline{\mathtt{v}}$ entails the inversion of $MV4V3$ and $V2$ thus resulting in $MV4V3V2$ of type $\mathtt{v}$,

and finally that $V1$ be a $(\mathtt{v}/\mathtt{v})$ and $MV4V3V2$ be v gives $V1MV4V3V2$ of type v.

7.2 Head Movement

Despite the promising view provided by remnant movement, we can still wish to keep so called *head movements*, particularly for traditional questions like V-to-I and I-to-C movements. In our previous (troncated) examples, we have not yet tried to give an account of the introduction of inflection features. We know that *inflection* is responsible for the nominative case. In our present words, that means that only an inflected sentence has an attractor for nominative case. In previous works ([18]), Stabler used to distinguish between weak and strong categorial features (for instance =V, =C, ... besides =v, =c, ...). If we follow the new trend, and particularly hints given by Chomsky, according to whom:

> The concept strength, introduced to force violation of Procrastinate, [also] has no place – which is just as well; strength was a feature of a feature, not an attractive notion

we are lead to use this notion of strength of a feature as rarely as possible. In fact, it seems that in most cases, the design of lexical entries (particularly the precise way to endow them with phonological and semantic features) enables us to dispense with the opposition weak/strong for features. It is an open question to know whether this distinction must be kept as "a true imperfection" or not.
From now on, we assume an inflection feature be introduced by means of a category $((\overline{\mathtt{k}}\backslash\mathtt{t})/\mathtt{vp})$, "phonologically" associated with some inflectional morpheme. The difference between an infinitival (*to read*) and a tensed verb (*reads*, *read*...) lies in the fact that the former can be directly inserted in a derivation (directly here means: without any prior hypothesis) while the latter is associated with a $\bullet$-product, like:

$$reads \quad ::= \quad \vdash ((\overline{\mathtt{k}}\backslash\mathtt{t})/\mathtt{vp}) \bullet ((\overline{\mathtt{k}}\backslash(\mathtt{d}\backslash\mathtt{vp}))/\mathtt{d})$$

It is worth to notice here that this provides an extension of our previous typing system. We assume from now on that types are no longer only of the form:

$$((F_2\backslash(F_3\backslash...(F_n\backslash(G_1 \otimes G_2 \otimes ... \otimes G_m \otimes A))))/F_1)$$

but they can be also $\bullet$-products of such types.
Otherwise, this leads us to re-use a rule that we had provisionally abandonned: the [$\bullet$ E] rule, which requires discharged hypotheses to be ordered and composed by ';'.

Rule:

$$\frac{\Gamma \vdash A \bullet B \quad \Delta[(A;B)] \vdash C}{\Delta[\Gamma] \vdash C}\,[\bullet E]$$

The reason to introduce this new device is linked to the HMC[10] which is assumed to still hold as soon as we keep this notion of Head Movement. It means that two head movements never cross (while they are allowed to cross phrasal movements): non commutativity of $\bullet$ rules out such a crossing and therefore "encapsulates" HMC.
As a matter of fact, because we know a head movement can cross a phrasal one, that can be too strong a constraint. It is therefore necessary to introduce a kind of *structural rule* which allows to perform associativity w.r.t. ';' and ',', a rule that we name *mixed associativity*, along the lines of [14], [11] etc.

Mixed Associativity rule:

$$\frac{\Theta[(\Gamma;(\Delta,\Delta'))] \vdash A}{\Theta[((\Gamma;\Delta),\Delta')] \vdash A}\,[MA]$$

We can then get the following deduction:

$$\dfrac{\vdash p : \overline{\mathtt{k}} \otimes \mathtt{d} \qquad \dfrac{\dfrac{y^2 : \overline{\mathtt{k}} \vdash y^2 : \overline{\mathtt{k}} \qquad \dfrac{\vdash likes : ((\overline{\mathtt{k}}\backslash\mathtt{t})/\mathtt{vp}) \bullet ((\overline{\mathtt{k}}\backslash(\mathtt{d}\backslash\mathtt{vp}))/\mathtt{d}) \quad subproof}{x^2 : \mathtt{d} \vdash likes\ x^2\ m\ likes\ m : (\overline{\mathtt{k}}\backslash\mathtt{t})}\,[\bullet E]}{y^2 : \overline{\mathtt{k}}; x^2 : \mathtt{d} \vdash y^2\ likes\ x^2\ m\ likes\ m : \mathtt{t}}\,[\backslash E]}{y^2 : \overline{\mathtt{k}}, x^2 : \mathtt{d} \vdash y^2\ likes\ x^2\ m\ likes\ m : \mathtt{t}}\,[entropy]}{\vdash p\ likes\ p\ m\ likes\ m : \mathtt{t}}\,[\otimes E]$$

[10] Head Movement Constraint

subproof:

$$\dfrac{y^1 : \overline{\mathtt{k}} \vdash y^1 : \overline{\mathtt{k}} \qquad \dfrac{\xi : ((\overline{\mathtt{k}}\backslash(\mathtt{d}\backslash\mathtt{vp}))/\mathtt{d}) \vdash \xi : ((\overline{\mathtt{k}}\backslash(\mathtt{d}\backslash\mathtt{vp}))/\mathtt{d}) \quad x^1 : \mathtt{d} \vdash x^1 : \mathtt{d}}{\xi : ((\overline{\mathtt{k}}\backslash(\mathtt{d}\backslash\mathtt{vp}))/\mathtt{d}); x^1 : \mathtt{d} \vdash \xi\ x^1 : (\overline{\mathtt{k}}\backslash(\mathtt{d}\backslash\mathtt{vp}))}[/E]}{\dfrac{y^1 : \overline{\mathtt{k}}; \xi : ((\overline{\mathtt{k}}\backslash(\mathtt{d}\backslash\mathtt{vp}))/\mathtt{d}); x^1 : \mathtt{d} \vdash y^1\ \xi\ x^1 : (\mathtt{d}\backslash\mathtt{vp})}{y^1 : \overline{\mathtt{k}}, \xi : ((\overline{\mathtt{k}}\backslash(\mathtt{d}\backslash\mathtt{vp}))/\mathtt{d}), x^1 : \mathtt{d} \vdash y^1\ \xi\ x^1 : (\mathtt{d}\backslash\mathtt{vp})}[entropy]}[\backslash E]$$

$$\dfrac{\vdash m : \overline{\mathtt{k}} \otimes \mathtt{d} \qquad y^1 : \overline{\mathtt{k}}, \xi : ((\overline{\mathtt{k}}\backslash(\mathtt{d}\backslash\mathtt{vp}))/\mathtt{d}), x^1 : \mathtt{d} \vdash y^1\ \xi\ x^1 : (\mathtt{d}\backslash\mathtt{vp})}{\xi : ((\overline{\mathtt{k}}\backslash(\mathtt{d}\backslash\mathtt{vp}))/\mathtt{d}) \vdash m\ \xi\ m : (\mathtt{d}\backslash\mathtt{vp})}[\otimes E]$$

$$\dfrac{x^2 : \mathtt{d} \vdash x^2 : \mathtt{d} \qquad \xi : ((\overline{\mathtt{k}}\backslash(\mathtt{d}\backslash\mathtt{vp}))/\mathtt{d}) \vdash m\ \xi\ m : (\mathtt{d}\backslash\mathtt{vp})}{x^2 : d; \xi : ((\overline{\mathtt{k}}\backslash(\mathtt{d}\backslash\mathtt{vp}))/\mathtt{d}) \vdash x^2\ m\ \xi\ m : \mathtt{vp}}[\backslash E]$$

$$\dfrac{\zeta : ((\overline{\mathtt{k}}\backslash\mathtt{t})/\mathtt{vp}) \vdash \zeta : ((\overline{\mathtt{k}}\backslash\mathtt{t})/\mathtt{vp}) \qquad x^2 : d; \xi : ((\overline{\mathtt{k}}\backslash(\mathtt{d}\backslash\mathtt{vp}))/\mathtt{d}) \vdash x^2\ m\ \xi\ m : \mathtt{vp}}{\dfrac{(\zeta : ((\overline{\mathtt{k}}\backslash\mathtt{t})/\mathtt{vp}); (\xi : ((\overline{\mathtt{k}}\backslash(\mathtt{d}\backslash\mathtt{vp}))/\mathtt{d}); x^2 : \mathtt{d})) \vdash \zeta x^2 m \xi m : (\overline{\mathtt{k}}\backslash\mathtt{t})}{\dfrac{(\zeta : ((\overline{\mathtt{k}}\backslash\mathtt{t})/\mathtt{vp}); (\xi : ((\overline{\mathtt{k}}\backslash(\mathtt{d}\backslash\mathtt{vp}))/\mathtt{d}), x^2 : \mathtt{d})) \vdash \zeta x^2 m \xi m : (\overline{\mathtt{k}}\backslash\mathtt{t})}{((\zeta : ((\overline{\mathtt{k}}\backslash\mathtt{t})/\mathtt{vp}); \xi : ((\overline{\mathtt{k}}\backslash(\mathtt{d}\backslash\mathtt{vp}))/\mathtt{d})), x^2 : \mathtt{d}) \vdash \zeta x^2 m \xi m : (\overline{\mathtt{k}}\backslash\mathtt{t})}[MA]}[entropy]}[/E]$$

7.3 Expletives and other phenomena

Let us look at the following paradigm, an example from Chomsky:

`(i) they [elected an unpopular candidate]`

`(ii) there was [elected an unpopular candidate]`

`(iii) an unpopular candidate was elected`

`(iv) there was an unpopular candidate elected`

Chomsky argues that "case (ii) involves feature attraction alone, yielding agreement of INFL and its associate 'candidate'. Case (iii) involves feature attraction followed by *Merge* of "an unpopular candidate" (=P(F')) to the category headed by INFL, yielding dislocation to SPEC-INFL. That some phrase must be merged in this position is determined by the Extended Projection Principle (EPP), an independant property. The choice of what is merged is determined by the initial numeration NUM; if NUM contains an expletive, that is merged, by preference of *Merge* over *Move*; if not, then the most local accessible phrase is merged – in (iii), the pied-piped phrase P(F')."
If this was true, the expletive "there" would be the bearer of a nake case-feature, attracted by $\overline{\mathtt{k}}$ in the subject position. We would have some particular syntactic object (perhaps built by using the [$\otimes$ I]-rule – not yet used –):

$$(\textit{there} : \overline{\mathtt{k}}) \otimes (\textit{an unpopular candidate} : \mathtt{d})$$

with the phonological feature *there* raising to the specifier position whereas *an unpopular candidate* remaining in place. There would be a rivality with the other, more classical, object:

$$\textit{an unpopular candidate} : \overline{\mathtt{k}} \otimes \mathtt{d}$$

Because *was elected* is supposed to be of type $((\overline{\mathtt{k}}\backslash\mathtt{vp})/\mathtt{d})$, we would obtain in the second case (that means (iii)):

an unpopular candidate was elected an unpopular candidate

instead of:

there was elected an unpopular candidate

Such a phenomenon could occur in many other contexts. We could have for instance:

`a man sleeps`

*there sleeps a man

Chomsky's arguments do not therefore give the good explanation. We suggest here that getting (ii) in preference to (iii) or vice-versa depends on the use of *be*. Suppose we have the following lexicon:

Lexicon:

$$\begin{array}{lcl} \mathit{there} & ::= & \vdash \mathit{there} : \overline{\mathtt{k}} \\ \mathit{was} & ::= & \vdash \mathit{was} : (\overline{\mathtt{k}}\backslash(\mathtt{t}/\mathtt{vp})) \\ \mathit{was} & ::= & \vdash \mathit{was} : ((\overline{\mathtt{k}}\backslash\mathtt{t})/(\overline{\mathtt{k}}\backslash\mathtt{vp})) \\ \mathit{elected} & ::= & \vdash \mathit{elected} : ((\overline{\mathtt{k}}\backslash\mathtt{vp})/\mathtt{d}) \end{array}$$

and suppose *an unpopular candidate* be of type $\overline{\mathtt{k}} \otimes \mathtt{d}$. We have the two following deductions[11].

$$\dfrac{\vdash \mathit{auc} : \overline{\mathtt{k}} \otimes \mathtt{d} \qquad \dfrac{\dfrac{\vdash \mathit{there} : \overline{\mathtt{k}} \quad \vdash \mathit{was} : (\overline{\mathtt{k}}\backslash(\mathtt{t}/\mathtt{vp}))}{\mathit{there\ was} : (\mathtt{t}/\mathtt{vp})}[\backslash E] \quad \dfrac{y : \overline{\mathtt{k}} \vdash y : \overline{\mathtt{k}} \quad \dfrac{\vdash \mathit{elected} : ((\overline{\mathtt{k}}\backslash\mathtt{vp})/\mathtt{d}) \quad x : \mathtt{d} \vdash x : \mathtt{d}}{x : \mathtt{d} \vdash \mathit{elected}\ x : (\overline{\mathtt{k}}\backslash\mathtt{vp})}[/E]}{y : \overline{\mathtt{k}}, x : \mathtt{d} \vdash y\ \mathit{elected}\ x : \mathtt{vp}}[\backslash E]}{y : \overline{\mathtt{k}}, x : \mathtt{d} \vdash \mathit{there\ was}\ y\ \mathit{elected}\ x : \mathtt{t}}[/E]}{\vdash \mathit{there\ was\ an\ unp\ cand\ elected\ an\ unp\ cand} : \mathtt{t}}[\otimes E]$$

$$\dfrac{\vdash \mathit{auc} : \overline{\mathtt{k}} \otimes \mathtt{d} \qquad \dfrac{y : \overline{\mathtt{k}} \vdash y : \overline{\mathtt{k}} \quad \dfrac{\vdash \mathit{was} : ((\overline{\mathtt{k}}\backslash\mathtt{t})/(\overline{\mathtt{k}}\backslash\mathtt{vp})) \quad \dfrac{\vdash \mathit{elected} : ((\overline{\mathtt{k}}\backslash\mathtt{vp})/\mathtt{d}) \quad x : \mathtt{d} \vdash x : \mathtt{d}}{x : \mathtt{d} \vdash \mathit{elected}\ x : (\overline{\mathtt{k}}\backslash\mathtt{vp})}[/E]}{x : \mathtt{d} \vdash \mathit{was\ elected}\ x : (\overline{\mathtt{k}}\backslash\mathtt{t})}[/E]}{y : \overline{\mathtt{k}}, x : \mathtt{d} \vdash y\ \mathit{was\ elected}\ x : \mathtt{t}}[\backslash E]}{\vdash \mathit{an\ unp\ cand\ was\ elected\ an\ unp\ cand} : \mathtt{t}}[\otimes E]$$

We must notice that the former proof can equally be presented as a deduction of (iv):*there was an unpopular candidate elected*. Let us also compare with French examples like:

(i) un candidat a été élu

(ii) il a été élu un candidat

(iii) un candidat a été élu

(iv) il y a eu un candidat élu

We can assume here that the locution *il y a* is already given the type $(\mathtt{t}/\mathtt{vp})$ in the lexicon.
Some remarks can be made after these examples:

- such a treatment of the expletive problem requires only one thing more: that we now accept *second order*-types in the lexicon,that means types which take as their arguments not only atomic features but also complex ones (like here $(\overline{\mathtt{k}}\backslash\mathtt{vp})$), something which is not possible in the earlier frameworks (like Stabler's minimalist grammars). This kind of generalization is due to the 'categorial grammarian' character of our calculus,
- going back to Chomsky's arguments: they are globally taken into account: when the expletive *there* 'belongs to NUM', it **must** be consumed, and that necessarily entails the selection of the verbal form $\mathit{was} : (\overline{\mathtt{k}}\backslash(\mathtt{t}/\mathtt{vp}))$, when it doesn't, the deduction is correct only if the 'auxiliary' attractor $\overline{\mathtt{k}}$ associated with the past participle is itself consumed by *was*, which at this time, is of type $((\overline{\mathtt{k}}\backslash\mathtt{t})/(\overline{\mathtt{k}}\backslash\mathtt{vp}))$,
- things do not result from a 'preference of *Merge* over *Move*' properly speaking, because there is a *Move* step in both analyses, but rather from the difference between two *Merge*-steps, one relying on a *second order-Merge*-step.

[11] where obvious entropy steps are omitted.

It is interesting to notice that Chomsky establishes a strong relation between these expletive constructions and 'long-distance effects' like they occur in unbounded dependencies. In fact, we can see that it is the very same mechanism as for the above expletives which can explain long-distance extractions like in:

(i) `what book do you think that Mary likes?`

(ii) `what book do you think Mary likes?`

(iii) `who do you think likes those books?`

if we use the following lexicon:

Lexicon

what book	::=	$\vdash \overline{\mathtt{wh}} \otimes \overline{\mathtt{k}} \otimes \mathtt{d}$
do	::=	$\vdash ((\overline{\mathtt{wh}} \backslash \mathtt{cp})/\mathtt{t}) \bullet ((\overline{\mathtt{k}} \backslash \mathtt{t})/\mathtt{vp})$
likes	::=	$\vdash ((\overline{\mathtt{k}} \backslash \mathtt{t})/\mathtt{vp}) \bullet ((\overline{\mathtt{k}} \backslash (\mathtt{d} \backslash \mathtt{vp}))/\mathtt{d})$
think	::=	$\vdash ((\mathtt{d} \backslash \mathtt{vp})/(\overline{\mathtt{wh}} \backslash \mathtt{cp}))$
think	::=	$\vdash ((\mathtt{d} \backslash \mathtt{vp})/\mathtt{t})$
that	::=	$\vdash ((\overline{\mathtt{wh}} \backslash \mathtt{cp})/\mathtt{t})$

Obviously, the occurrence of a three-components product (and by generalizing, of an n-components product) compels us to slightly change the [$\otimes$ E] and [$\bullet$ E] rules, and our 'second criterion on admissible proofs'. The criterion is now formulated as:

Proposition 3 *(***(new) second criterion on admissible proofs***) n hypotheses ($n \geq 2$) are discharged by a product π_n of n factors according to the order First In, First Out, or by a product π_{n+1} of n+1 factors, strictly containing π_n, in which case, the n+1 hypotheses are discharged in such a way that the last hypothesis introduced in the deduction (which corresponds to the $(n+1)^{th}$ factor) is discharged immediately after introduction, and are therefore also said to be discharged in the order First In, First Out.*

The formulation of the new [$\otimes$ E] is the following, (that of [$\bullet$ E] can be deduced from it straightforwardly).

$$\frac{\Gamma \vdash \{\alpha, \epsilon, ..., \epsilon\} : A_n \otimes A_{n-1} \otimes ... \otimes A_1 \quad \Delta, x_n : A_n, x_{n-1} : A_{n-1}, ..., x_1 : A_1 \vdash \gamma : C}{\Gamma, \Delta \vdash \gamma[\{\alpha, \epsilon, ..., \epsilon\}/\{x_1, x_2, ..., x_n\}] : C} [\otimes E]$$

where some hypotheses A_i may be starred and where α is substituted to the x_i labelling the most recent hypothesis ${A_i}^*$, or if there is not, the first A_i introduced (which is generally labbelled A_1).

8 Conclusion

We have presented here a logical ground for 'minimalist' ideas, by following Chomsky's proposal as close as we could. This is not simply for sake of logics... Following Husserl and Cavaillès [2], we believe in a *mathesis universalis*, expressed by formal logic, the aim of which is to reconstruct any object *in general* (*"Que fait le mathématicien sinon décrire ou fixer ce qui concerne tout objet, élément abstrait d'une multiplicité?"*[2, p.48]). What is convincing in Chomsky's work is precisely this effort towards exploring the linguistic object *in general*, starting from minimal hypotheses. We may postulate that such an enterprise, which is rather *speculative* in fact, is on the lines of logic, as it was conceived by such philosophers. The proposal we make, of applying systems which were created from a more general perspective to linguistics, is therefore meaningful in the sense that in doing so, we describe a linguistic object as a particularisation of more general ones, thus trying to accomplish a step on the way towards the *unification of science*.
Such an application is also legitimated by the insights it provides on new solutions concerning some linguistic problems like expletives and unbounded dependencies.
It remains to show how theoretical results on such logical systems can be applied to the particular one in order for instance to introduce new insights on parsability and computational complexity. It seems already possible to claim that polynomiality is ensured by the fact that partially non-commutative linear logic, limited to elimination rules is already polynomial (C. Retoré, forthcoming).

References

[1] M. Abrusci and P. Ruet. Commutativity and non-commutativity. Technical report, Universita di Roma Tre, Rome, 1998.

[2] Jean Cavaillès. *Sur la logique et la théorie de la science*. Bibliothèque de Philosophie Contemporaine. Presses Universitaires de France, Paris, 1960.

[3] N. Chomsky. *The Minimalist Program*. The MIT Press, 1996.

[4] N. Chomsky. Minimalist inquiries: the framework. Technical report, MIT, Cambridge, 1998.

[5] P. de Groote, editor. *The Curry-Howard isomorphism*. Cahiers du Centre de Logique. Université catholique de Louvain-la-Neuve, Louvain-la-Neuve, 1995.

[6] P. de Groote. Partially commutative linear logic: sequent calculus and phase semantics. In M. Abrusci and C. Casadio, editors, *Proofs and Linguistic Categories*, pages 199–208. CLUEB - University of Chieti, 1996.

[7] M. Dymetman. Group theory and computational linguistics. *Journal of Logic, Language and Information*, 7:461–497, 1998.

[8] J.Y. Girard. Linear logic. *Theoretical Computer Science*, 50(1):1–102, 1987.

[9] R. Kayne. *The Antisymmetry of Syntax*. The MIT Press, 1994.

[10] H. Koopman and A. Szabolcsi. Verbal complexes. Technical report, UCLA, Los Angeles, forthcoming.

[11] N. Kurtonina and M. Moortgat. Structural control. In P. Blackburn and M. de Rijke, editors, *Specifying Syntactic Structures*, Studies in Logic, Language and Information, pages 75–113. CSLI Publications, Stanford, 1997.

[12] J. Lambek. The mathematics of sentence structure. *American Mathematical Monthly*, 65:154–170, 1958.

[13] A. Lecomte and C. Retoré. Towards a minimal logic for minimalism. In G. van Kruijf, editor, *Formal Grammar'99*, Utrecht, 1999. ESSLLI'99.

[14] M. Moortgat. Categorial type logics. In J. van Benthem and A. ter Meulen, editors, *Handbook of Logic and Language*, chapter 2, pages 93–178. Elsevier, 1997.

[15] G. Morrill. *Type Logical Grammar, Categorial Logic of Signs*. Kluwer, Dordrecht, 1994.

[16] A. Ranta. *Type Theoretical Grammar*. Oxford University Press, 1994.

[17] C. Retoré. *Calcul des séquents ordonnés*. Thèse, Universite Paris-7, 1993.

[18] E. Stabler. Derivational minimalism. In C. Retoré, editor, *Logical Aspects of Computational Linguistics*, volume 1328 of *LNCS/LNAI*, pages 68–95. Springer, 1997.

[19] E. Stabler. Remnant movement and complexity. Technical report, UCLA, Los Angeles, 1999.