# Structure Dependence and Linear Order: Clarifications and Foundations

Jordi Fortuny[1]

[1] *Departament de Filologia Catalana, Facultat de Filologia, Universitat de Barcelona, Gran Via de les Corts Catalanes, 585. Barcelona 08007, Spain*

September 24, 2014

**Abstract**

According to Chomsky (2010, 2013) and Berwick et al. (2011), the Structure Dependence Principle suggests that linear order is a reflex of the sensory-motor system and plays no role in syntax and semantics. However, when these authors use the expression *linear order* they seem to refer exclusively to the literal precedence relation among terminals in linguistic objects. We observe that this narrow use differs from the technical use in a non-innocuous way and does not allow us to exploit the unificational force that the concept of order can have for minimalist investigations. Here we follow Fortuny & Corominas' (2009) formal definition of the syntactic procedure that capitalizes on the foundational set-theoretical concept of *nest*. We show how the Structure Dependence Principle can be derived from a local definition of syntactic domain while retaining the idea that central concepts of configurational and transformational syntactic theories are orders.

# Contents

# 1 The Structure Dependence Principle

In this paper we shall focus on a plausible constraint on possible grammars, the so-called Structure Dependence Principle (1), postulated by Chomsky (1972), which can be informally expressed in the following terms:

(1) *The Structure Dependence Principle* (SDP)
Syntactic operations are defined on hierarchical representations

Consider, for concreteness, the following well-known auxiliary fronting case study, where the interrogative sentence (2-b) is formed from the declarative sentence (2-a).

(2) a. Peter is tall
b. Is Peter tall?

The paradigm given in (2-a) and (2-b) is consistent with the following three rules, since the interrogative is formed by commuting two words of the declarative sentence (3), or more particularly by fronting the first occurrence of *is* (4), which is also the the second word of the declarative sentence (5):

(3) Given a declarative sentence form the appropriate interrogative sentence by commuting any two contiguous words

(4) Given a declarative sentence form the appropriate interrogative sentence by fronting the first occurrence of “is” in the declarative sentence

(5) Given a declarative sentence form the appropriate interrogative sentence by fronting the second word in the declarative sentence

Each of these three rules makes different predictions about the formation of an interrogative sentence from the novel declarative sentence (6), as illustrated, respectively, by the three strings (6-a), (6-b) and (6-c).

(6) The man who is talking is tall
  a. *The man who is talking tall is?
  b. *Is the man who talking is tall?
  c. *Man the who is talking is tall?

However, none of these strings is grammatical, whereby none of the three rules can be generalized adequately to form the grammatical question (7).

(7) Is the man who is talking tall?

The three rules under discussion share the property of being expressed in terms of precedence/temporal relationships among words; however, in order to attain the correct grammatical rule, we must bring into consideration the hierarchical structure of sentences. The approximative hierachical representation of (6) is provided in (8).

(8) [ [the man [who is talking] ] is tall]

We can now informally express the following rule, which is intuitively structure-dependent and accounts satisfactorily for the question-formation of both (2-a) and (6).

(9) Given a declarative sentence form the appropriate question by fronting the hierarchically most prominent auxiliary in the declarative sentence

This leads us to the generally accepted conclusion that in order to determine the appropriate grammatical system with the correct predictions it is necessary to consider the hierarchical properties of linguistic expressions, instead of the position of words in the precedence relation. Therefore, the metrics for movement operations is structurally determined.

## 1.1 A clarification

We must consider in more detail the connection between the literal precedence/temporal relationship among the terminals of an expression and the SDP, informally expressed in (1).

Let us begin by noting that we cannot properly conclude from the need to claim that syntactic operations are structure dependent, and not dependent on the precedence relation, that *linear order* plays no role in syntax and semantics but is only relevant to the sensory-motor system. This seems to be the view contended by Chomsky (2010, 2013) and Berwick et al. (2011). For instance, Chomsky (2010) claims that "the best explanation for the choice of structural rather than linear distance would be that linear order is simply not available to the operations of the I-language –that it is a secondary phenomenon imposed by the sensory motor system (p. 11)", and Berwick et al. (2011) state that "linear order seems to be a reflex of the sensory-motor system, and so unavailable to the syntax and semantics we describe there (p. 1217)". According to this position, the syntactic and semantic components have *hierarchy* and *structure*, but not *linear order*.

In the view we have just sketched (and more generally, in generative grammar) the use of the expressions *order*, *linear* and *linear order* differs from their technical meaning: they refer exclusively to the precedence/temporal relationship among terminals, which is indeed an order. However, it is clear that there are order relations that are not precedence/temporal relations. We emphasize this elementary but important terminological issue, because, as will become clear, it is not merely a particular informal use that innocuously differs from the technical use, but rather the source of obscurity about an aspect of the theory of grammar that has a prominent place in current generative linguistics, namely how abstract hierarchical representations are assigned a precedence/temporal relationship. This aspect is usually called (again misleadingly) *the linearization problem*.

What is essential of order relations (in the technical sense) is that they are transitive and antisymmetric. By definition, a relation $R$ is:

(10) a. transitive iff, for any objects $x$, $y$, $z$, if $\langle x, y\rangle \in R$ and $\langle y, z\rangle \in R$,

then $\langle x, z \rangle \in R$, and

b. antisymmetric iff, for any two objects $x$, $y$, if $\langle x, y \rangle \in R$ and $\langle y, x \rangle \in R$, then $x = y$.

Let us consider, for clarity, some typical order relations. The "greater than" ($>$) and "the greater than or equal" ($\geq$) relations defined in the set $\mathbb{N}$ of natural numbers are linear orders in $\mathbb{N}$; they are orders because they are transitive and antisymmetric, and they are linear in $\mathbb{N}$ because they are connected in $\mathbb{N}$. In general,

(11) A relation $R$ is connected in $A$ iff, for any two objects $x, y \in A$, either $\langle x, y \rangle \in R$ or $\langle y, x \rangle \in R$.

The terms "linear" and "connected" are thus synonimous when they refer to order relations. The difference between $>$ and $\geq$ is that the former is a strict linear order in $\mathbb{N}$ and the latter is a reflexive linear order in $\mathbb{N}$. As usual,

(12) a. A relation $R$ is strict (or irreflexive) iff, for all object $x \in A$, $\langle x, x \rangle \notin R$.
b. A relation $R$ is reflexive iff, for all object $x \in A$, $\langle x, x \rangle \in R$.

Other instances of orders are the strict inclusion relationship ($\subset$) and its reflexive associate ($\subseteq$). We can observe that they are non-linear in the set of subsets of $\mathbb{N}$, since we can find sets of natural number such as $\{1, 2\}$ and $\{7, 8, 9\}$ that are not connected by these relations. The two pairs of relations that we have briefly considered ($>/\geq$, $\subset/\subseteq$) are orders but none of them involves a precedence/temporal relationship.

# 2 Nests

We shall take as a starting point Fortuny & Corominas' (2009) formalization of certain core aspects of current generative syntax (cfr., among others, Stabler 2011 for a different perspective). This work is a particular technical implementation of the minimalistic approach to phrase structure initiated by Chomsky (1995), which capitalizes on the intuition that the simplest way to generate hierarchically structured expressions is by means of a successive operation that takes two objects as input and merges them together, thereby yielding as output a new object composed solely of the two objects taken as input. Chomsky's earliest minimalist considerations also argued for the methodological virtue of dispensing with strictly grammatical idiosyncracies (such as X'-theory, or trace-theory). A further though narrowly related possibility was indicated by Epstein (1999) in his study of c-command: namely, that syntactic relationships could simply derive from the computational procedure, they would not constitute primitive syntactic elements but rather computational by-products (cfr. also Epstein 1998 for a broader derivational study of syntactic relations). Fortuny & Corominas' note is a particular execution of the reductionist program initiated by Chomsky and Epstein: it provides a precise definition of both the computational procedure and the basic hierarchical notions of the generated outcomes (constituents, chains of copies, syntactic domain, dominance relationship) that does not rely on idiosyncratic grammatical elements but solely on the most fundamental elements of set theory. This proposal also confirms Kayne's (1994) influential intuition that an ordering among the terminals of a linguistic expression can be directly obtained from hierarchical syntactic relations, although X'-theory does not need to be assumed. Fortuny's (2008) original idea that nests may offer a useful tools for theoretical syntax has also been adapted by Zwart's (2009, 2010) work on syntactic dependencies and de Belder & van Craenenbroeck's (2011) proposal on root insertion.

## 2.1 Simple nesting

Let us thus apply the formalization developed in Fortuny & Corominas 2009 to study in a precise way the generation of expression (7), repeated in (13).

(13) Is the man who is talking tall?

Let the alphabet for the generation of (13) be the set $A$ of minimal syntactic categories, each viewed as a singleton whose single element is a primitive element, also called a terminal,

$$A = \{\{the\}, \{tall\}, \{talking\}, \{man\}, \{who\}, \{is\}\}.$$

We now consider how the nesting machine $\mathcal{N}$ proceeds in order to generate (13). At the first step $s_0$, $\mathcal{N}$ selects an element $\{k\}$ of a given alphabet and generates the set $\{k_0\}$, which we refer to as the set $M_0$. Accordingly, we shall say that when an element $\{k\}$ of an alphabet comes into the computation at step $s_0$, its element, $k$, becomes the occurrence $k_0$. Assume for concreteness that in the derivation of (13) $\mathcal{N}$ generates the singleton $\{tall_0\}$, i.e.:

$$s_o : \{tall_0\} = M_0.$$

At a derivational step $s_{n>0}$, $\mathcal{N}$ selects an element $\{k\}$ of an alphabet, forms the set $\{k_n\}$ (namely the set whose element is the $n$-occurrence of $k$) and generates the set $M_n$ which is recursively defined as the union of $\{k_n\}$ and the output $M_{n-1}$ of the immediately previous step $s_{n-1}$. Assume $\mathcal{N}$ continues the derivation started in $s_0$ with the following union-formation operation:

$$s_1 : M_0 \cup \{is_1\} = \{tall_0\} \cup \{is_1\} = \{tall_0, is_1\} = M_1.$$

$\mathcal{N}$ is thus recursively defined as an operation that generates at the first step $s_0$ the set $M_0 = \{a_0\}$ for $\{a\}$ an element of a given alphabet $A$, and forms, at any further step $s_{n>0}$, the union of the output $M_{n-1}$ and $\{k_n\}$ for $\{k\} \in A$. When an arbitrary element of $A$, namely $\{e\}$, comes into the computation at step $s_{i\geq 0}$, its element, $e$, becomes an occurrence $e_i$.[1] This procedure can be symbolically

[1]Cfr. Chomsky 1955 for the need for a notation to distinguish among occurrences. Chomsky's notation was based on a devise proposed by Quine (1940) and differs from Fortuny & Corominas' (2009) proposal, which distinguishes occurrences by keeping track of the derivational step where they are introduced.

summarized as follows:

$$
\begin{aligned}
&M_0 = \{a_0\}\,(\{a\} \in A)\\
&M_{n>0} = M_{n-1} \cup \{k_n\}\,(\{k\} \in A),
\end{aligned}
$$

where $n$ is unboundedly large but finite; this means that a nesting computation must always be finite although there is no fixed natural number that constrains the length for nesting computations.

Before continuing with the study of the generation of expression (13), it will be useful to introduce a precise definition of the notion of *constituent*, which plays a crucial role in configurational approaches to syntax in determining how semantic relations are set on the basis of syntactic representations.

> **Definition of constituent**. A constituent $C_k$ is the outcome of $\mathcal{N}$ at a particular step $s_k$, which is the set of those sets $M_j$ $(0 \leq j \leq k)$ successively generated,
>
> $$C_k = \{M_0, M_1, ..., M_k\}\,.$$

> **Remark 1**. Any constituent $C_k$ generated by $\mathcal{N}$ is a *nest*, i.e., a set of sets linearly ordered by inclusion,
>
> $$M_0 \subset M_1 \subset ... \subset M_k.$$

Let us proof by induction that the outcome $C_n$ resulting from stopping $\mathcal{N}$ at step $s_{n\geq 0}$ is always a nest (Remark 1), which is precisely the same thing as saying that, for any two distinct elements $M_i, M_j \in C_n$, either $M_i \subset M_j$, or $M_j \subset M_i$. Consider firstly that we stop $\mathcal{N}$ at $s_0 : \{a_0\}$; accordingly, the outcome of this derivation will be, by the Definition of constituent, the constituent

$$C_0 = \{\{a_0\}\}\,,$$

which is a trivial nest, since it is vacuously true that all distinct elements of $C_0$ are pairwise related by inclusion. Consider now that we stop $\mathcal{N}$ at an arbitrary

$s_n$ when $n > 1$ and assume, by inductive hypothesis, that $C_{n-1}$ is a nest, i.e., $M_0 \subset ... \subset M_{n-1}$; since $M_n$ is the union of $M_{n-1}$ and a singleton, we conclude that $M_{n-1} \subset M_n$. Accordingly, $M_0 \subset ... \subset M_{n-1} \subset M_n$, whereby $C_n$ is also a nest when $n > 1$. Therefore, any constituent is a nest.

**Remark 2**. Given a nesting derivation of $n$-steps, the set $\mathcal{C}_n$ of constituents generated during this derivation,

$$\mathcal{C}_n = \{C_i : i \leq n\},$$

is a nest (i.e., it is a set of sets linearly ordered by inclusion).

We can determine, for instance, that in the nesting derivation we are studying there are so far two constituents, one for each step:

$$\begin{aligned} C_0 &= \{M_0\} = \{\{tall_0\}\} \\ C_1 &= \{M_0, M_1\} = \{\{tall_0\}, \{tall_0, is_1\}\}. \end{aligned}$$

The set $\mathcal{C}_1$ of constituents generated during this derivation is

$$\mathcal{C}_1 = \{\{\{tall_0\}\}, \{\{tall_0\}, \{tall_0, is_1\}\}\},$$

which is indeed linearly ordered by inclusion.

Therefore, the proposed formal definitions of constituent and of set of constituents of a syntactic object are nests.

A further interesting consequence of introducing nests into syntactic theory is that the precedence/temporal ordering among the terminals of an expression is straightforwardly obtained from the hierarchical or nested representation. As proved by kuratowski (1921), a family $F$ of sets linearly orders a set $S$ iff $F$ is saturated as to the property of being a nest of $S$ (i.e., iff all the elements of $F$ are subsets of $S$ and $F$ is not a proper subset of a nest of $S$). For instance, the family $F_1 = \{\{a\}, \{a, b\}, \{a, b, c\}\}$ is a linear order of the set $S = \{a, b, c\}$, since all the elements of $F_1$ are subsets of $S$ and there is no nest of $S$ that is a proper superset of $F_1$. But the families $F_2 = \{\{a\}, \{a, b, c\}\}$ or $F_3 = \{\{a, b\}, \{a, b, c\}\}$, which are also nests of $S$, are not linear orders of $S$, given that they both are subsets of $F_1$.[2]

> **Remark 3**. For a given constituent $C_n$, we can identify the set $M_n$ as the set of occurrences of syntactic terminals; note that $C_n$ is always saturated as to the property of being a nest of its element $M_n$, whereby $C_n$ is a linear order of $M_n$. In other words, a constituent is a linear order of the occurrences of its terminals. This linear ordering can be straightforwardly interpreted as the precedence/temporal relation among the occurrences of syntactic terminals of a constituent at the sensory-motor system.

Given the constituent

$$C_1 = \{\{tall_0\}, \{tall_0, is_1\}\},$$

the set of terminals is

$$M_1 = \{tall_0, is_1\}.$$

Since $C_1$ is a nest saturated of $M_1$, it can be interpreted as the ordering

$$< is_1, tall_0 > .$$

In sum, the nesting machine generates syntactic constituents, which are linear orders on the basis of which the precedence/temporal relation can be directly read of at the sensory-motor system.

It is now crucial to observe that the next element to be introduced in the nesting derivation of (13) is not a terminal, but a complex syntactic object:

(14) [the [man [who [is [talking]]]]]

[2] Cfr. Hallett 1986 for a detailed discussion of Kuratowski's method of reducing the notion of *order* to set theory and its precedents.

## 2.2 Complex nesting

Both on empirical and conceptual grounds it is necessary to allow syntactic computations to manipulate not only minimal syntactic categories but also complex syntactic objects (cfr. Zwart 2009 for a broader discussion on layered derivations). With this purpose Fortuny and Corominas allow the nesting machine to be compounded of multiple derivational spaces, labelled as $D_1$, $D_2$, ..., $D_n$, and define necessary conditions as well as the required notation.

The two conditions imposed on nesting computations involving multiple derivational spaces are the following:

1. the number of derivationals spaces involved in a nesting derivation must be bounded
2. at the end of a given derivation, only one derivational space can remain open, which implies that all derivational spaces used in the derivation but one generate their outcomes as inputs for other derivational spaces.

As for notation, subindexes and superindexes refer, respectively, to derivational steps and derivational spaces. For instance, the terminal $k_f^g$ would have been introduced at the derivational space $D_g$ and at the step $s_f$, the set $M_r^d$ would have been formed at $D_d$ and at $s_r$, the constituent $C_i^j$ would have been generated at $D_j$ at step $s_i$, and the set $\mathcal{C}_i^j$ would contain all those constituents generated a $D_j$ at some $s_{f \leq i}$. When the constituent $C_s^i$ is the final outcome of $D_i$, and $D_i$ feeds $D_j$ at step $s_r$, then $C_s^i$ becomes $C_r^{i/j}$. In order to keep our representations as simple as possible, we will introduce indexes only when they are required for practical reasons, i.e., only when we need to distinguish different occurrences of the same element.

Let $M_1^1 = \left\{tall, is_1^1\right\}$ and assume that the constituent depicted in (14) is generated at $D_2$:

**$\mathbf{D}_2$**

$$\begin{array}{lcl} s_0 & : & \{talking\} \\ s_1 & : & \left\{talking, is_1^2\right\} \\ s_2 & : & \left\{talking, is_1^2, who\right\} \\ s_3 & : & \left\{talking, is_1^2, who, man\right\} \\ s_4 & : & \left\{talking, is_1^2, who, man, the\right\} \end{array}$$

$$\begin{aligned} C_4^2 \quad = \{ \quad & \{talking\}, \left\{talking, is_1^2\right\}, \left\{talking, is_1^2, who\right\}, \left\{talking, is_1^2, who, man\right\}, \\ & \left\{talking, is_1^2, who, man, the\right\} \ \} \end{aligned}$$

At this point $D_2$ must feed $D_1$, which means that the final outcome $C_4^2$ of $D_2$ is introduced into $D_1$. Accordingly, $\mathcal{N}$ takes as input $\left\{C_4^2\right\}$, forms the set $\left\{C_5^{2/1}\right\}$ and performs the operation $\left\{tall_0^1, is_1^1\right\} \cup \left\{C_5^{2/1}\right\}$ thereby generating the new constituent $C_5^1$. See Table 1 for a representation of the whole derivation involving two spaces of $C_5^1$.

Table 1: Nesting derivation of $C_5^1$. Derivational space $D_2$ consists of five derivational steps; its final outcome is the constituent $C_4^2$. In $D_1$ at $s_5$, $\mathcal{N}$ takes as input $C_4^2$, forms the set $\left\{C_5^{2/1}\right\}$, that is the final outcome of $D_2$ reintroduced at $D_1$ at $s_5$, and performs the operation $\left\{tall_0^1, is_1^1\right\} \cup \left\{C_5^{2/1}\right\}$. The outcome of $D_1$ is the set $C_5^1$.

| | $D_2$ | $D_1$ |
|---|---|---|
| $s_0$ | $M_0^2 = \{talking\}$ | $M_0^1 = \{tall\}$ |
| $s_1$ | $M_1^2 = \{is_1^2, talking\}$ | $M_1^1 = \{is_1^1, tall\}$ |
| $s_2$ | $M_2^2 = \{who, is_1^2, talking\}$ | |
| $s_3$ | $M_3^2 = \{man, who, is_1^2, talking\}$ | |
| $s_4$ | $M_4^2 = \{the, man, who, is_1^2, talking\}$ | |
| | $C_4^2 = \{M_0^2, M_1^2, M_2^2, M_3^2, M_4^2\}$ | |
| $s_5$ | $\Longrightarrow$ | $M_5^1 = \{C_5^{2/1}, is_1^1, tall\}$ |
| | | $C_5^1 = \{M_0^1, M_1^1, M_5^1\}$ |

Since $C_5^1$ is the linear order

$$< C_5^{2/1}, is, tall >$$

and $C_5^{2/1}$ the linear order

$$< the, man, who, is_1^2, talking >,$$

the following precedence/temporal ordering among terminals is obtained by composition of relations:

$$< the, man, who, is^2_1, talking, is^1_1, tall > .$$

It is not necessary to introduce any complication to obtain an ordering among the terminals of a linguistic expression; X'-theory does not need to be assumed and Kayne's (1994) so-called Linear Correspondance Axiom becomes a mere byproduct of the generative procedure, as argued in Fortuny 2008. It is also remarkable that the problems posed to Kayne's approach by non-branching complements and by branching specifiers do not appear once we dispense with X'-notation and we bring into consideration appropriate set-theoretical tools, as illustrated.[3]

[3] In principle, a linear order such as

$$C = \{\{John\}, \{John, kisses\}, \{John, kisses, Mary\}\}$$

could be read as a precedence relationship or as a successor relationship, respectively,

$$< John, kisses, Mary >$$

$$< Mary, kisses, John > .$$

If $C$ is interpreted as a successor relationship, then $\mathcal{N}$ is top-down procedure that first merges the specifier of a projection and forms the trivial constituent $C_0$ containing only the specifier, then introduces the head and forms the constituent $C_1$ containing solely the speficier and the head, and finally introduces the object and forms the final constituent $C_2$ containing the speficier, the head and the object.

$$\begin{array}{lll} s_0 & : & \{John\} \\ & C_0 & = \{\{John\}\} \\ s_1 & : & \{John, kisses\} \\ & C_1 & = \{\{John\}, \{John, kisses\}\} \\ s_2 & : & \{John, kisses, Mary\} \\ & C_2 & = \{\{John\}, \{John, kisses\}, \{John, kissesMary\}\} \end{array}$$

Accordingly, the constituents we obtain if $\mathcal{N}$ is read as a successor relationship are in contradiction with standard constituency considerations. However, if $\mathcal{N}$ is assumed to be a top-down procedure, then we obtain the correct expected constituency (cfr. Fortuny & Corominas 2009). Note that the top-down procedure just defined differs from Phillips' (2003) left-to-right procedure for generating syntactic structures, which creates a new constituent by destroying a previous one. Indeed, Phillips' procedure is not top-down, but rather left-to-right.

## 2.3 Internal-merge and external merge

We shall contend, adapting Chomsky's (2001, 2008) terminology, that $X$ is "externally merged" to $Y$ when $X$ is either a minimal category selected from the alphabet or a complex syntactic object generated at a different derivational space, and "internally merged" to $Y$ when it is selected from the syntactic domain of $Y$. We emphasize, adapting Chomsky's (2001:n. 29) view, that the virtue of allowing $\mathcal{N}$ to perform internal merge operations is a matter of applicability and of conceptual necessity: it is a matter of applicability because it is a simple analytical tool that is constructively used to capture the property of "displacement", which seems ubiquitous in natural languages; and a matter of conceptual necessity because only by stipulation could $\mathcal{N}$ be banned to perform internal merge operations, a stipulation that would seem unmotivated, given our current understanding of the syntactic patterns of natural languages.

Recall that, as noted in section 1, we form a question by fronting the "hierarchically most prominent" occurrence of an auxiliary (9), and not by fronting the first occurrence of an auxiliary in the precedence/temporal relationship (4). In other words, locality conditions on movement operations are structure dependent, they are not defined on the basis of the precedence/temporal relationship. The definition of (syntactic) domain must ensure that the occurrence $is_1^1$ can be fronted to form the appropriate question, whereas the occurrence $is_1^2$ cannot. We thus provide the following local definition of domain.

> **Local definition of domain**. Given a set $M_j^i$, $\{x\}$ belongs to the domain of $M_j^i$ ($\Delta(M_j^i)$) iff one of the following conditions is fulfilled: (a) $x$ is a constituent generated at $D_i$ at some step previous to $s_j$ (or in other words, $x$ belongs to the set $\mathcal{C}_{j-1}^i$ of constituents), or (b) $x$ is an element of $M_j^i$. Symbolically,
>
> $$\Delta(M_j^i) = \left\{\{x\} : x \in \mathcal{C}_{j-1}^i \vee x \in M_j^i\right\}.$$

Accordingly, the domain of $M_5^1 = \left\{tall, is_1^1, C_5^{2/1}\right\}$ is the set

$$\Delta(M_5^1) = \mathcal{C}_{5-1}^1 \cup M_5^1 = \left\{C_0^1, C_1^1, C_2^1, C_3^1, C_4^1, tall, is_1^1, C_5^{2/1}\right\}.$$

Given that

$$\{is_1^1\} \in \Delta(M_5^1)$$

and that

$$\{is_1^2\} \notin \Delta(M_5^1),$$

we can perform at $s_6$ of $D_1$ the operation (1) but not the operation (2).

$$M_5^1 \cup \{is_1^1\} \tag{1}$$

$$M_5^1 \cup \{is_1^2\} \tag{2}$$

This accounts for the grammaticality contrast between (13) and (6-b), as desired.

Note that, if the syntactic domain of $M_5^1$ were the linear order

$$C_5^1 =< the, man, who, is_1^2, talking, is_1^1, tall >,$$

we would expect $is_1^2$ to be a legitimate target for a fronting operation, if not the favorite one, given that it is closer to the landing position than $is_1^1$. The empirical observation behind the SDP reveals that the syntactic domain of a set $M_5^1$ cannot be the linear order $C_5^1$, but the set $\Delta(M_5^1)$ . The SDP does not indicate that linear order is irrelevant for the syntactic component whatsoever, but rather that the domain for internal merge operations is not a linear order of occurrences but a set of syntactic categories defined in such a way that it bans internal merge operations to cross derivational spaces.[4] If this remark is neglected, then we prevent ourselves from grounding several basic syntactic notions on a single general concept, namely that of order.

We conclude this study on the relationship between linear order and structure dependence by observing that we can define the chains of copies generated by

[4]We refer the interested reader to Zwart 2009 for a broader study of opacity effects in terms of layered derivations, i.e., in terms of derivations involving multiple derivational spaces, using our terminology. Zwart (2009, 2010) argues as well for the derivation of the LCA from a syntactic algorithm that generates nests and studies how the notion of *dependency* can be defined from an ordered syntactic representation. The procedure described therein differs from ours in being top-down (cfr. also Fortuny & Corominas 2009).

internal merge operations as nests, i. e., as linear orders (cfr. Nunes 2004). Whereas we say that when the element $x$ is externally merged at $s_j$ it becomes an occurrence $x_j$, we shall say that it becomes a copy $x_{j/i}$ when it is selected from $s_j$ and remerged at $s_{i>j}$. We thus define the notion of *chain* in the following terms:

> **Definition of chain**. A chain $CH(x_j)$ is a linear order of the copies of an occurrence $x_j$,
>
> $$CH(x_j) = \left\{ \left\{ x_{j/k}, ..., x_j \right\}, \left\{ ..., x_j \right\}, \left\{ x_j \right\} \right\}.$$
>
> The copy $x_j$ is the tail of $CH(x_j)$, the copy $x_{j/k}$ the head and any $x_{j/n(j<n<k)}$ an intermediate copy. Multiple copies of $x$ are identified by virtue of the subindex referring to the derivational step where the occurrence has been introduced to the derivation, and distinguished with respect to each other by virtue of the subindexical suffix referring to the derivational step where they are subsequently merged $({}_{/i})$. For a given constituent $C_i^j$, $M_i^j$ is the set of occurences and copies involved in $D_j$. The head of the chain is pronounced by the sensory-motor system, whereas the tail and usually the intermediate copies are not externalized, although they remain active at the conceptual-intentional system.

Since $is_1^1 \in \Delta(M_5^1)$ and $is_1^2 \notin \Delta(M_5^1)$, $\mathcal{N}$ can generate the outcome (3) but not the outcome (4)

$$C_5^1 = \left\{ \left\{ tall \right\}, \left\{ tall, is_1^1 \right\}, \left\{ tall, is_1^1, C_2^{2/1} \right\}, \left\{ tall, is_1^1, C_5^{2/1}, is_{1/3}^1 \right\} \right\} \tag{3}$$

$$C_5^1 = \left\{ \left\{ tall \right\}, \left\{ tall, is_1^1 \right\}, \left\{ tall, is_1^1, C_5^{2/1} \right\}, \left\{ tall, is_1^1, C_5^{2/1}, is_{1/3}^2 \right\} \right\} \tag{4}$$

The chain of the occurrence $is_1^1$ formed by internal merge at $D_1$ is:

$$CH(is_1^1) = \left\{ \left\{ is_1^1 \right\}, \left\{ is_1^1, is_{1/3}^1 \right\} \right\},$$

where $is_1^1$ and $is_{1/3}^1$ are respectively the tail and the head.

Note finally that this structure dependent rule also accounts for the observation that (15) is unambiguously the question corresponding to (15-a) (Berwick et al. 2011): the occurrence of the auxiliary *can* generated in the matrix clause in (15-a) is hierarchically more prominent than the one generated in the relative clause in (15-b). The former is externally merged in the main derivational space $D_1$, whereas the latter is externally merged in a secondary derivational space $D_2$ that feeds $D_1$. As a consequence, the former auxiliary, but not the latter, belongs to the search domain, and thus can be internally merged in $D_1$ in order to form the question in (15), which is, consequently, unambiguous.

(15) Can eagles that swim fly?
  a. Eagles that fly can eat
  b. Eagles that can fly eat

# 3 Conclusion

In the previous section we have applied Fortuny & Corominas' (2009) theory of hierarchical expressions constructed on the basis of the concept of *nest* to the study of a well-known condition on syntactic operations: the Structure Dependence Principle (SDP). The observation behind this principle reveals that syntactic operations in natural languages are not carried out by scanning the literal precedence relationship among the terminals of a sentence but by taking into account the structures into which they enter (Chomsky 1972).

We have accounted for the structure-dependence effects in terms of a local definition of syntactic domain that reduces computational complexity by prohibiting internal merge operations to cross derivational spaces. As a consequence, the syntactic domain for internal merge operations is not given by a precedence relationship among terminals; however, this does not entail that linear order plays no role in syntax. Indeed, the concepts of *constituent* and *chain* can be defined on the basis of nests, the foundational set-theoretical notion to which order can be reduced (Kuratowski 1921); it is also worth noting, in this regard, that a constituent is a linear order of occurrences and copies, whereby the literal precedence/temporal order relation among terminals can be directly derived from our definition of constituent. As these considerations reveal, the technical notion of order can have a remarkable unificational force for theoretical linguistics. In this sense, it is not the case that order is "secondary phenomenon" only relevant to the sensory-motor system, but it is an essential property of syntax and of the syntactic representations that feed the conceptual-intentional system.

Under the light of our proposal, the SDP does not seem to be a primitive or unmotivated principle of UG, contrary to the view originally expressed in Chomsky (1972):

> "the structure dependent operation has no advantages from the point of view of communicating efficiency or 'simplicity'. If we were, let us say, designing a language for formal manipulations by a computer, we would certainly prefer structure independent operations. These are far simpler to carry out, since it is only necessary to scan the words of sentences paying no attention to the structure into which they enter, structures that are not marked physically in the sentence at all (p. 28)".

Firstly, the claim that "structures are not marked physically in the sentence at all" is too strong, since, as shown in several experimental works, prosody plays an auxiliary role in indicating how words are combined into phrases (cfr., among many others, Prieto 1997). Thus, structures are partially marked in speech.

Secondly, structure independent operations based solely on the relation of literal precedence among "words" are clearly insufficient for setting grammatical operations: they would fail to account for the manipulation of a complex constituent, as in the internal merge of a whole constituent, for which it is necessary to bring into account the structure into which words enter; in brief and artificial language that preferred structure independent operations would not be able to generate a sentence like *the man who is tall is talking*, whose subject is an internally organized syntactic unit. Therefore, structure dependence, or the capacity of manipulating linearly ordered structures, has a very clear advantage at least in terms of syntactic and semantic richness.

And thirdly, the observation that internal merge operations are structure dependent can be related to simplicity considerations: derivations involving multiple spaces may be required to ensure a certain level of syntactic and semantic richness, but internal merge operations crossing derivational spaces are problematic. The structure dependence of internal merge operations is, thus, advantageous from the point of view of structural or derivational "simplicity".

Our technical developments have thus an interesting consequence concerning the content and the nature of UG: the SDP, which seemed to be an idiosyncratic element of UG that could not be motivated on independent grounds, can now be viewed as a natural consequence of how the generative procedure of Language forms constituents and chains by minimizing the search space.

We believe thus that the methodological virtue of our theory of hierarchical expressions contributes substantially to a deeper understanding of the nature of the syntactic component of Universal Grammar. Last but not least, as noted in Fortuny & Corominas 2009, the concept of nest is a powerful abstract entity postulated in different domains, like Theoretical Biology, Statistical Physics or Genetics (see Fortuny & Corominas 2009 for relevant references), where a recursive algorithm or an evolutionary process are involved.

## Acknowledgements

This work has been supported by the Secretary for Universities and Research of the Ministry of Economy and Knowledge of the Government of Catalonia and the Cofund programme of the Marie Curie Actions of the 7th R&D Framework Programme of the European Union, the research projects 2009SGR1079, and FFI201123356. I would like to thank the audience of the miniseminar I gave at the University of Cambridge in 2013 where I first presented the main idea of this note; I am particularly indebted to Ian Roberts, Theresa Biberauer, Michelle Sheehan and Jenneke van der Wall for being willing to hear about these issues and for the multiple questions they raised. I have also benefited from presenting this work in the Incontro di Grammatica Generativa held in Trento in 2014. I am especially grateful to Jan-Wouter Zwart and Bernat Corominas, for their collaboration and for their contributions to the nesting machine, and finally to Adriana Fasanella for carefully reading a previous version of this note and providing several improvements.

# References

De Belder, M., & J. van Craenenbroeck. How to merge a root. Ms., CRISSP Brussels.

Berwick, R. P. Pietroski, B. Yamkama & N. Chomsky. 2011. Poverty of stimulus revisited. *Cognitive Science* 35(7):1207–124.

Chomsky, N. 2013. Problems of projection. *Lingua* 130:33–49.

Chomsky, N. 2010. Poverty of stimulus: Unfinished business. Lecture presented in the Lecture Series 'Sprache und Gehirn-Zur Sprachfahigkeit des Menschen', organized by A. D. Friederici in the context of the Johannes Gutenberg endowed professorship.

Chomsky, N. 2008. On phases. In *Foundational Issues in Linguistic Theory*, ed. In R. Freidin, C. P. Otero & M. L. Zubizarreta, 133–166. Cambridge, MA: MIT Press.

Chomsky, N. 2001. Beyond explanatory adequacy. *MIT Occasional Working Papers in Linguistics* 20:1-18. Cambridge, MA: MIT Press.

Chomsky, N. 1995. Categories and transformations. *The minimalist program.* Cambridge, MA: MIT Press.

Chomsky, N. 1972. *Problems of knowledge and freedom.* The Russell lectures, Vintage Books.

Chomsky, N. 1955/1975. *The logical structure of linguistic theory.* Plenum Press.

Epstein, S. 1999. Un-principled syntax. In *Working minimalism*, ed. S. Epstein & N. Hornstein, 317–345. Cambridge, MA: MIT Press.

Epstein, S. 1998. *A derivational approach to syntactic relations.* Oxford: Oxford University Press.

Fortuny, J. & B. Corominas. 2009. Some formal considerations on the generation of hierarchically structured expression. *Catalan Journal of Linguistics* 8:99–111.

Fortuny, J. 2008. *The emergence of order in syntax.* Linguistik Aktuell/Linguistics Today. John Benjamins Publishing Company.

Hallet, M. 1986. *Cantorian set theory and limitation size.* Oxford: Oxford University Press.

Kayne, R. 1994. *The antisymmetry of syntax.* Cambridge, MA: MIT Press.

Kuratowski, K. 1921. Sur la notion de l'ordre dans la théorie des ensembles. *Fundamenta mathematicae* 2:161–171.

Nunes, J. 2004. *Linearization of chains and sideward movement.* Cambridge, MA: MIT Press.

Prieto, P. (1997). Prosodic Manifestation of Syntactic Structure in Catalan. In *Issues in the Phonology of the Iberian Languages*, ed. F. MartÃnez-Gil & A. Morales-Front, 179–199. Washington D.C.: Georgetown University Press.

Quine, W. 1940. *Mathematical logic.* Cambridge, MA: Harvard University Press.

Stabler, E. P. 2011. *Computational perspectives on minimalism.* In Oxford handbook of linguistic minimalism, ed. Cedric Boeckx, 617–643. Oxford: Oxford University Press.

Zwart, J-W. 2009. Prospects for top-down derivations. *Catalan Journal of Linguistics* 8:197-187.

Zwart, J-W. 2010. Structure and order: asymmetric merge. In *Oxford Handbook of Linguistic Minimalism*, ed. C. Boeckx, 96–119. Oxford: Oxford University Press.