1. Introduction. Language and Mind

As known, the first half of the 20th century is dominated by structuralism (Manoliu-Manea 1974); since linguistic structuralism largely focused on the study of phonological and morphological systems, structuralists soon came to the conclusion that languages vary in unpredictable ways and without assignable limits (Joos, 1957), so that linguistic theory could at best be a specification of general methods of dealing with the data, methods which were viewed as discovery procedures for possible grammars of given languages (Harris 1959). With very limited exceptions (Wells 1949), syntax was not formally studied, either because it represented the use, not the invariant structure of language (i.e. “parole”, not “langue” in de Saussure’s terms), or because the methods used in phonology and morphology could not very well be extended to syntax.

The interest for discovering forms (shapes) and patterns in syntax is historically linked to the use in linguistics of the notion of universal grammar (=UG), while the particular conception of UG which is entertained is inevitably linked to the definition of language that one adopts. The second half of the 20th century witnesses two crucial proposals for UG. The first proposal is due to Noam Chomsky, who, in a series of works spanning several decades, has founded the still mainstream generative enterprise. Chomsky dares to use the term UG only as late as his 1965 book, though he had elaborated the first form of his proposal ten years earlier, in the monumental Logical Structure of a Linguistic Theory. The second proposal is Richard Montague’s 1968 paper, explicitly called Universal Grammar, which represented the culmination of a long tradition of approaching languages with the tools of formal logic.

Despite the skeptical attitude that they had for each other’s work, Montague and Chomsky share important common goals, such as: constructing explicit syntactic and semantic representations of sentences, defining the operations necessary to construct these representation, building these representations from “bottom to top”, through procedures explicitly relating the lexicon to syntax.

The differences between them are also considerable and stem from the different definition of language that they adopt. For Montague (1970/1974: 222), “there is no important theoretical difference between natural languages and the artificial languages of logicians”, since the syntax and semantics of both can be comprehended “within a single natural and mathematically precise theories.” Languages are mathematical entities, specifically algebras with relations defined on them. Montague is interested in uncovering the logical form of sentences; logical forms serve a double purpose. First, these representations, in conjunction with the meaning of certain items, namely, those endowed with logical meaning (e.g. modal operators, quantifiers), should define the class of valid inferences generated by each sentence. Secondly, these logical forms can and must be interpreted, the ultimate goal being that of constructing a theory of truth for a natural language fragment.
Chomsky’s interests are vastly different: for him natural languages are psychological entities, so that linguistics is part of cognitive psychology and should be looked upon as a natural, rather than social science. According to Chomsky, this conclusion is immediately apparent to anyone who seriously considers the theoretical significance of linguistic creativity and language acquisition. Linguistic creativity is the speaker’s ability to produce and understand an infinite number of sentences, many of which he has not uttered or heard before. For instance, confronted with Kaplan’s surprising title “How to Russell a Frege-Church”, speakers have no difficulty to understand that in this case the proper name Russell is used as a verb, so that the title probably means “how to give a Russelian solution to a Frege-Church problem”. Since people have finite memories, it is plausible to assume that they may produce and understand an infinite number of sentences because they have internalized a grammar, a finite device which provides a semantic and phonological interpretation for any sentence of some given language L. Any grammar of L will project the finite and somewhat accidental set of observed utterances to a presumably infinite set of grammatical utterances the speaker can produce and understand.

The second empirical fact that generative theory calls attention is that grammars and thus languages are learnable. By the age of five, any normal child has already acquired his grammatical competence; he will have mastered the morpho-syntactic structures of his language; in later years, he will mostly enrich his language at the level of the vocabulary. It should be obvious that grammar is acquired without explicit instruction, despite the complexity of the task. It is highly unlikely that infants are ever instructed on how to form relative clauses, on word order differences, or on the significance of particles like still, yet, etc. The inevitable conclusion is that grammars are acquired because children are genetically endowed with what a Descartes first called a faculté de langage, actually a language acquisition device which helps the child process the linguistic input he gets and construct his internalized grammar. The FL as well as the internalized grammar are components of mind, whence the claim that linguistics is part of cognitive psychology.

According to Chomsky (1986), the proper domain of study of linguistics is thus the internalized language (I-lg) as opposed to the externalized language (i.e. E-lg, language as part of the environment), which acts as the stimulus that triggers off the development of the internalized language on the basis of the child’s innate faculty of language. Chomsky’s notion of language then is that of I-lg, defined as what the speaker knows which enables him to speak and understand a language. The linguist should consequently focus on three tasks: 1) defining what we know when we know a language; 2) accounting for how language is acquired; 3) accounting for how language is put to use.

Putting it more formally, I-lg is defined as the “system of knowledge attained and internally represented in the mind/brain. [Chomsky 1986: 20]”. I-lg is the content of the so-called language Steady State, S_S, attained by the language faculty in an adult speaker through a process of maturation. The language state S_S is a relatively steady state, which may undergo only peripheral changes, such as the loss or addition of lexical concepts. The system of knowledge which is the content of I-lg is arrived at on the basis of the child’s linguistic experience (i.e. exposure to E-lg), starting from the initial state S_0 of the language faculty. As to the content of S_0, the faculty of language at birth, it clearly may be viewed as a Universal Grammar: “UG is construed as a system of conditions deriving from the human biological endowment that identifies the languages that are humanly accessible under normal conditions [Chomsky 1986: 24]”. These universal conditions limit the possible variation of natural languages and define human languages as a natural kind.
**Conclusion:** Chomsky defines the faculty of languages in its mature state as an *I-lg*, as the finite system of knowledge possessed by the speaker and represented on his mind. FL is part of the biological endowment, a module of mind, like vision, hearing.

2. **From Rule Systems to Principles and Parameters**

Within the generative paradigm, there have been several implementations of this view of *I-lg* (Chomsky 1965, 1981, 1995) differing with respect to their descriptive and explanatory power, as well as regarding the *type of universals that they entail* and the role played by the foundational axiom that UG is innate and *I-lg* is a component of mind/brain.

2.1. **The Standard Theory**

The first definition of an *I-lg*, adopted in the *Standard Theory* (Chomsky 1965, 1955/1973) is that of a grammar as a **finite system of rules** generating all and only the correct sentences of a language.

Sentence derivations work from top to bottom. Sentences are assigned an analysis by means of phrase structure rules, which give the possible constituents of *phrases* (i.e. S, NP, VP, etc.). Phrases are eventually broken down into *lexical categories* (i.e. units of the lexicon like N, V, etc.), replaced by suitable lexical items. Here is a derivation of sentence (1a). Notice that given the top to bottom direction of the analysis, lexical items play no special role.

(1) a. \[ ([\text{Tom}_{NP}] \, [\text{[kissed}_V] \, [\text{a white monkey}_{NP} \, V_P])_S] \]

b. \[ S \rightarrow NP \land VP \]

\[ NP \rightarrow N \]

\[ VP \rightarrow V \land NP \]

\[ NP \rightarrow D \land A \land N \]

\[ N \rightarrow \text{Tom} \]

\[ V \rightarrow \text{kissed} \]

\[ D \rightarrow a \]

\[ A \rightarrow \text{white} \]

\[ N \rightarrow \text{monkey} \]

(2) a. \[ [\text{The white monkey}_{NP}] \, [\text{was}_AUX] \, \text{kissed by Tom}. \]

b. \[ [\text{Was}_AUX] \, [\text{the white monkey}] \, \text{kissed by Tom}?. \]

c. \[ \text{Last night} \, [\text{the white monkey}_{NP}] \, [\text{was}_AUX] \, \text{kissed by Tom}. \]

d. \[ *[\text{Was}_AUX] \, \text{last night} \, [\text{the white monkey}_{NP}] \, \text{kissed by Tom}?. \]

(3)

```
  S
 / \  /
NP   VP
 /\  /\  /
N V   NP
 /\ /\ /\ /
Det A N
Tom kissed the white monkey
```

The main virtue of the analysis is that it assigns the sentence its *proper constituent structure*, showing which strings are *constituents* in the sentence, and what kind of constituents
they are; for instance, *kissed the* is not a constituent, while *a white monkey* is a constituent, moreover, an NP constituent. Constituent structure is the most important syntactic property of sentences, since all syntactic operations are *structure dependent*. For instance *Inversion* in English is an operation that applies in questions to an NP and an auxiliary verb in, provided that the NP is the subject and that it is in sentence initial position. These conditions are met in (2a), but not in (2c), which is why sentence (2d), where the auxiliary is fronted just as in (2b) is ungrammatical. The assigned constituent structure also shows that sentences, and generally syntactic units, are *hierarchical*, not linear. The hierarchical arrangement of constituents may be shown in a labeled bracketing (1a), or in a *tree diagram* like (3). In the standard theory, trees merely sum up the information already supplied by phrase structure rules, regarding constituency, dominance and order.

This conception of I-logic envisages formal, rather than substantive, universals. All languages have the *same types of rules* (phrase structure rules and transformational rules). Rules are *construction specific* inside one language and may be *language specific* cross-linguistically. For instance, Romanian adjectives are normally post-posed, so that a noun phrase like *o maimuță albă* is the output of a different phrase structure rule: NP $\rightarrow$ D^N^A.

We retain that the rule-systems to approach I-logic has no *principled way of handling variation* and grants a modest role to language universals.

**Principles and Parameters Grammars**

The picture will change radically under the second generative conception of I-logic, the *principles and parameters [=P&P] grammar*. The mature human language faculty is now described as a set of *principles*, shared by all languages, together with *a set of parameters*, which *follow from the principles and determine the possible variations manifest in languages*. In this conception, variation is *principled and limited*, since it follows from the *finite number of parameters of UG*. The P&P grammar is thus the first empirically successful attempt to cope with the Tower of Babel problem. Secondly, in this conception, UG moves from merely representing formal universals to (also) representing *substantive ones*: the principles give properties true of all languages representing *substantive universals*. In contrast, the parameters express the cross-linguistic variability of languages. An easy example is the *Extended Projection Principle*, which says that all sentences have subjects. However, according to the *Null Subject Parameter* only in certain languages the subject is always *overtly expressed* (it is phonological content); in other language, the subject may be inaudible, being represented as a phonologically null pronoun (*pro*).

(4) **UG Principle**: *The Extended Projection Principle: All sentences have subjects*

**Parameter**: *The Null Subject Parameter*

In some languages the subject must be overtly expressed (English, French), in some others it may be phonologically null (Romanian).

(5) a. *pro* citea un roman.

b. He was reading a novel.

**Aim of the paper**

In this paper we turn to another property of P&Ps grammars: the use of *configurational concepts*. The P&P grammar exploits to a greater extent the fact that *I-logs* are components of
mind. One important consequence of this fact is that grammatical information need not be represented only by verbal, propositional knowledge, but visual images, *shapes* may also be stored and relevant in producing and understanding sentences. Specifically we briefly examine the concept of *c-command* and the *phrase structure template*, both with a major role in structuring syntax.

3. Endocentricity and the Phrase Structure Template

An important characteristic of P&P grammars is the re-*lexicalization of syntax*. Since there are no rules, syntax may only rely on the indisputable fact that knowing a language primarily means knowing the words of the language. The assumption is that through their formal properties (i.e. syntactic category and selectional properties), words encapsulate their grammar since they implicitly specify admissible modes of combination. Knowing words means knowing how to use them in phrases and sentences, which amounts to knowing the grammar of the language. The direction of analysis is reversed; sentence construction starts from the *lexicon*, building syntactic representation *from bottom to top*. Lexical items are *heads* which generate phrases.

In this context, an important result was the discovery that all phrases (i.e. NPs, APs, etc.) not only turn round *heads* but also have an *isomorphic structure*. The *head* of the phrase is the lexical item which “attracts” the other constituents of the phrase. For instance a preposition like *to* in (6) minimally attracts a NP and maximally attracts some other constituent, for instance, a nominal measure phrase like *three meters*. Similarly, a transitive verb like *cut* obligatorily attracts an NP as its object, it normally also attracts an NP subject (7).

(6)  a.  *He moved [toP].
   b.  He moved [[toP] [the left NP] P]PP]
   c.  He moved [[three meters NP] [toP] [the left NP] P]PP].

(7)  a.  *[Cutting V] was nice.
   b.  [ [Cutting V] [the cake NP] V] VP was nice.
   c.  [[John NP] [was] [cutting V] [the cake NP] V] VP

As already suggested by these examples, phrases have the same configuration irrespective of the syntactic category of the head, i.e. regardless of whether the head is a noun (N), a verb (V), an adjective (A) or a preposition (P). Lexical heads *project phrases* attracting obligatory and optional constituents. The constituent which is obligatory for the *minimal phrase of a particular type* is called the *complement of the head*. Selection of the complement is *syntactic*, since heads select complements of a designated syntactic category. For instance, in English, verbs select NPs (*break a window*) or PPs (*depend on friends*), while adjectives and nouns select PP complements (*aware/awareness of the danger/*aware the danger*). The first *projection X’ of some head X* is constituted of just the *head and its complement*. Outside the first projection, heads combine with *specifiers* on a semantic basis.

Summing up, there is a principle of UG, responsible for the projection of phrases, more exactly for the hierarchical arrangement of constituents inside a phrase. The principle is known as the *Principle of Endocentricity* (stated in (8)). On its basis, the *typical configuration* of a phrase in UG, known as the *phrase structure template* is as in (9). The examples given below
show that verb phrases, adjectival phrases, prepositional phrases and nominal phrases may be identical, differing only through the syntactic category of their complements.

(8) **Principle of endocentricity**
   a. Any word X (head) projects a number of supercategories, ending in a maximal projection, XP.
   b. Any phrase XP has a head X.

*The phrase structure template*

(9) \[
\begin{array}{c}
\text{XP} \\
\text{Specifier} \\
X' \\
X \quad \text{Complement} \\
\text{Head}
\end{array}
\]

\(X = \text{V, N, A, P}\)

(10) *John loves cats.*

\[
\begin{array}{c}
\text{VP} \\
\text{Specifier} \\
[\text{John}\text{NP}] \\
V' \\
V \quad \text{Complement} \\
\text{love} \\
[\text{cats}\text{NP}]
\end{array}
\]

(11) *[John’s fondness of cats] NP](impressed me).

\[
\begin{array}{c}
\text{NP} \\
\text{Specifier} \\
[\text{John’s}\text{NP}] \\
\text{fondness} \\
\text{Complement} \\
[\text{of cats}\text{PP}]
\end{array}
\]

(12) *(I consider)[John fond of cats] AP]*
The phrase structure template is a recurrent cognitive configuration, with an enormous descriptive power. Importantly, the template in (9) specifies only the hierarchy of constituents, the fact that the complement is inside the first projection of the head, while the specifier is in a higher domain. Just as the Extended Projection Principle, the Principle of Edocentricity leaves some room for variation. The parameter deriving from the fact that the phrase structure template does not specify word order is the Head Parameter. The main provision of the Head Parameter concerns the position of the head with respect to its complement.

(14) The Head Parameter

A head may precede its complement (head initial languages) or a head may follow its complement (head final languages).

English and modern Romance languages are head initial languages, for all the major complement-taking categories. Languages are often non-uniform regarding the head parameter. For instance German and Dutch are head final regarding verbal projections, but head-initial regarding the other lexical projections. Yet German and Dutch are considered head-final languages since the properties of the verb are decisive in characterizing a languages according to the head parameter.

(15) English

V read a book
P for a friend
A proud of his son
N destruction of the city

(16) Verbal projections, head final

German ein Buch lesen
Dutch een boek lezen

Prepositional projections, head initial

German mit einem Hammer (with a hammer)
Dutch met Marie (with Mary)

Results

The discovery of this invariant pattern, which is part of UG, eliminates the need for phrase structure rules and represented a considerable simplification of syntax. The task of the learner is also made considerably easier: he simply needs to acquire the value of the head parameter and then extend his vocabulary. A lexical item contains its grammar being responsible for the projection of a phrase.
3. **A configurational concept: C-command**

As already seen, syntactic hierarchies may be represented as trees, which define two types of relations, dominance (containment) and ordering relations. Both may alternatively be represented by labeled bracketing, as done above in (5) and (6).

Some concepts however seem to be irreducibly configurational and have only been defined on tree structures. This is the case of the concept of c-command (= constituent command, Reinhart 1983). C-command expresses the fact that \( \alpha \) and \( \beta \) are *in the same constituent*, but \( \alpha \) is higher, “more prominent”, than \( \beta \) and has certain priorities with respect to \( \beta \). The shape described by c-command is (10). Notice that the exact position of \( \beta \) with respect to \( \alpha \) is not only unspecified, but it is unspecifiable since the path relating the two constituents is a *variable*. To that extent, it is only the relative position of \( \alpha \) and \( \beta \) (= the shape) that matters.

\[
\text{C-command}
\]

(17) \( \alpha \) c-commands \( \beta \) iff the first branching node above \( \alpha \) is also above \( \beta \):

\[
\begin{array}{c}
X \\
\alpha \\
\beta \\
\end{array}
\]

C-command is a defining property of large number of linguistic processes. Among the processes which depend on *asymmetric c-command* we mention anaphoric relations, movement (which is always to a c-commanding position, Pesetsky (2012), linearization (Kayne, 1994) a.o. At this point, we give just one illustration:

**Relevance of c-command**

(18) Pronouns should not be c-commanded by their antecedents

a. Rosa\(_i\) complained that she\(_{ij}\) had a headache.

b. She\(_i\) complained that Rosa\(_{ij}\) had a headache.

c.  

\[
\begin{array}{c}
\text{TP} \\
\text{Rosa}_i \\
\text{She}_i \\
\text{T} \\
\text{VP} \\
\text{V} \\
\text{CP} \\
\text{complained} \\
\text{C} \\
\text{that} \\
\text{TP} \\
\text{she}_{ij} \quad \text{had a headache} \\
\text{Rosa}_{ij} \quad \text{had a headache} \\
\end{array}
\]

Sentence (18a) is interpretable under the given coindexation since the pronoun *she* is c-commanded by the antecedent *Rosa*, but does not c-command it. By contrast, sentence (18b)
cannot mean that Rosa complained that she, Rosa, head a headache, since this time the pronoun she c-commands the antecedent Rosa against the principle in (18).

5. A constraint on tree shapes. Reducing complex configurations to simpler ones

Originally, trees were very “free”, so that a node in a tree could have any number of branches. This uncontrolled complexity represented a potential problem for learnability. In particular when a verb had more than one complement, ternary branching was freely used, as in (22a), for a sentence like Mary gave John a book. In this section we argue that there are empirical reasons for rejecting the hypothesis of ternary-, and more generally, multiple-branching. Correct configurations are always binary (Kayne 1984) and heads cannot have more than one complement. This has the desirable result of a drastic simplification and reduction in the number of syntactic shapes admissible in natural languages.

Binary Branching. The Single Complement Hypothesis Let us review one argument that ternary branching is not a correct representation of verbs that have more than one nominal complement, such as give, offer, etc. The claim is that nominal arguments may not symmetrically c-command each other, so that configuration (22a), which involves ternary branching and allows the verb to have two complements on the same level is illicit. Configuration (22a) should be replaced by (22b) which observes the single complement hypothesis, in that each V head has exactly one complement and (possibly) one specifier

(19) The Single Complement Hypothesis (Larson 1988)
Each argument of a verb is introduced in a different verb shell.

More precisely, in (22b), the higher V head (a copy of the lower V) takes the subject, NP₁ as its specifier and takes one complement, namely, a VP. In the second VP, the second argument, NP₂ is projected as a specifier, and the third argument, NP₃ is projected as a single complement. NP₁ asymmetrically c-commands NP₂, while NP₂ asymmetrically c-commands NP₃.

Configurations (22a) and (22b) make different predictions regarding many linguistic phenomena (Larson 1988), among which the distribution of the negative polarity item any. As shown in (21), anyone is licensed only in (21b), where it is a complement c-commanded by the negative verb.

(20) Licensing of Polarity Item ANY
Any is licit only if it is (asymmetrically) c-commanded by overt negation:

(21) a. *Anyone didn’t see Mary
b. Mary didn’t see anyone.

(22) a.  
```
  VP
   NP₁  V'
      V  NP₂  NP₃
      Mary gave John a book
```
It is immediately apparent that configuration (22a, 23) makes the wrong predictions regarding the licensing of *any. Since NP$_2$ c-commands NP$_3$, one can indeed say *John gave no one anything. However since NP$_3$ also c-commands NP$_2$, one should also be able to say *John gave anyone nothing, contrary to fact. In contrast, in (22b, 24), only NP$_2$ c-commands NP$_3$, so that only the correct sentence *John gave no one anything is ruled in. Configuration (24), unlike (23) is no longer ambiguous with respect to c-command.

Thus syntax has adopted both the Single Complement Hypothesis and the constraint on binary branch.

The same analysis can be extended to other predicates that appear to have two complements, such as the class of verbs which select ___ NP + PP complements (*put, place, etc); they will project VPs like (25b). Through this process of lexical decomposition, complex three-argument configurations are reduced to standard two-argument configurations, eliminating symmetric c-command.

(25)  a. John put the books on the table
6. Conclusions
1. Configurational concepts like *the phrase structure template* and *c-command* are part of UG.
2. Syntax is responsible for the existence of a number of tree shapes, in terms of which sentences are generated and interpreted.
3. Lexical meanings are largely flexible and depend on the pattern where the word is inserted. This is one source of well-known plasticity of word meaning, as illustrated below:

(26)  
\begin{itemize}
  \item a. John walks every day.
  \item b. John walks his dog every day.
  \item c. John walked his way to a slimmer self this year.
  \item d. John has walked his shoes ragged.
\end{itemize}

References