

THE GENERATIVE ENTERPRISE¹

José Borges Neto²
(UFPR/CNPq)

In 1982, Riny Huybregts and Henk van Riemsdijk published a long interview with Noam Chomsky, organized as a discussion of different issues in the history and structure of the linguistic theory. The book containing that interview was given an extremely adequate title: ‘*The Generative Enterprise*’ (see Chomsky 1982a). I share the interviewees’ opinion that the linguistic theory, proposed, developed and permanently ‘protected’ by Chomsky’s extraordinary presence can in fact be understood as an ‘enterprise’ – a collective one - having Chomsky as its unequivocal leader.

My aim here is to demonstrate that the linguistic theory widely known as *Generative Grammar* (among other names it has been given at different times) is an extremely coherent Scientific Research Program that starts to be built around the mid XXth century and that has become, from its early days, a way of understanding the structure of the human language that can be questioned – an appropriate procedure in the development of scientific theories – but that cannot be ignored.

I shall consider the roughly 50 years of Generative Grammar (hereafter GG) as a period of time in which a *Scientific Research Program* is constructed, with roughly the same meaning the term has in the proposal presented by the Hungarian philosopher Imre Lakatos for approaching the history of sciences in general (see Lakatos 1978). Although I will not adopt an orthodox lakatosian approach, I believe Lakatos’s ideas allow for a better quality understanding of what actually happened to GG during the last 50 years.

Therefore, the present work is not a strictly linguistic work. Better, it is a work that fits the areas of History and Philosophy of Science³, directed to the analysis of a specific case: the ‘internal’ history of GG.

1. Lakatos’s methodology.

I shall start the presentation of Lakatos’s methodology quoting Feyerabend:

Let me now present in its entirety the picture of science which I think should replace Kuhn’s account.

This picture is the synthesis of the following two discoveries. First, it contains Popper’s discovery that science is advanced by a critical discussion of alternative views. Secondly, it contains Kuhn’s discovery of the function of tenacity which he has expresses, mistakenly I think, by postulating tenacious periods. The synthesis consists in Lakatos’ assertion ... that proliferation and tenacity do not belong to successive periods of the history of science, but are always copresent.

(Feyerabend 1974 p. 211)

¹ Published in Portuguese as “O empreendimento gerativo” In Mussalim, F. & Bentes, A.C. (eds) 2004. *Introdução à Lingüística*, vol 3. São Paulo: Cortez.

² I would like to thank my colleagues Maria Cristina de Figueiredo Silva and Evani Viotti. for the comments, ever pertinent, they made on a first version of the present work; and also the editors Fernanda Mussalim and Anna Christina Bentes for their welcome suggestions and comments. Of course, the responsibility for the final results is mine.

³ As Lakatos (1971:102) says, paraphrasing Kant: ‘*Philosophy of Science without history of science is empty; history of science without philosophy of science is blind*’.

Let us start with the concepts of proliferation and tenacity.

In Lakatos, *Proliferation* means that it is desirable to have theories in competition. According to him, the history of science has been – and so it should be – a history of scientific research programs in competition.

Tenacity, in Lakatos, means that the scientist does not abandon a theory that has been refuted, as Popper would like it to happen but, on the contrary, he does all he can to maintain it either ignoring the counter-examples or reanalyzing them in such a way as to transform them into validating evidences of the theory. In his own words:

*Nature may shout **no**, but human ingenuity – contrary to Weyl and Popper – may always be able to shout louder. With sufficient resourcefulness and some luck, any theory can be defended “progressively” for a long time, even it is false*
(Lakatos 1971, p. 111)

According to Lakatos, the best way of starting out the ‘science game’ is not with a falsifiable hypothesis, but with a *Scientific Research Program* (SRP) consisting of basically a *nucleus* and a *heuristics*.

The nucleus of a SRP is a set of propositions that, based on methodological decisions, are taken as ‘untestable’, that is, propositions that are sometimes said to be ‘metaphysical’ and that reveal the point of view that will orient the approach to the subject, the very definition of the object of study, etc.

The heuristics of a SRP is a set of methodological rules that tells us which directions must be taken in searching for scientific ‘explanations’. The heuristics is a kind of ‘development policy’ of the program, that is, a selection and ordering of problems, a plan that leads to a progressive sophistication of the explanatory models. It is a plan that establishes a sequence of reality simulation models, ever more and more complex, deep, and comprehensive.

Lakatos offers as an example of the action of such heuristics the developing process of the Newtonian program. Newton initially developed a model for a planetary system having only one planet gravitating around the sun in which both the sun and the planet were referred to as *points*. In that model, he managed to obtain the inverse square law for Kepler’s ellipse. The third law of dynamics, however, forbade such extremely simple model and Newton replaced it by another model in which both the sun and the planet rotated around the gravity center of a system formed by both. Next, Newton adapted the model in order to include more planets, admitting, however only heliocentric forces, but not interplanetary forces. Subsequently, he worked on the possibility of the sun and the planets being *spheres* and not points. This stage of development of the program demanded overcoming enormous mathematical difficulties. Problems solved, Newton started to work with *rotating spheres* and their oscillations. He accepted the interplanetary forces and started to work with perturbations. Later on, he worked with irregular planets instead of spherical planets, coming closer and closer to the real planetary systems⁴.

For Lakatos, then, the program progresses through the development of a series of models that result in *creative shifts* in the heuristics, in other words, in revisions made in the program’s ‘developing plan’.

⁴ For a more detailed presentation of the present case, see Lakatos 1970, p. 50-51.

2. GG: a Scientific Research Program.

The idea I have defended for some time now is that GG is an SRP and not a linguistic ‘theory’. I am not going to justify this position here (see Borges Neto 1991 for details). On this line of thought, instead of discussing alternative proposals developed by such and such linguists, it is more interesting to follow the major directions the program assumes, through its rules and heuristics. It is also interesting to see the creative shifts that from time to time redirect the efforts made by the scientists linked to the SRP.

Such is the investigation we intend do carry out within the scope of the present work.

2.1. The chomskian SRP: nucleus and heuristics.

Without further discussion, I would like to propose that the nucleus of the GG consists of the following statements:

- (1) The effective linguistic behavior (utterances) is, at least partially, determined by states of the mind/brain.
- (2) The nature of the states of the mind /brain, partially responsible for the linguistic behavior, can be captured by the computing systems that form and modify representations⁵.

I believe these two statements adequately summarize the concept of human language that has presided the fifty years of chomskian thought.

The GG program’s heuristics determines that the linguist’s basic task is to create computing systems that may serve as model for the speaker/hearers’ linguistic knowledge of the language. These computing systems must be understood as explanatory hypotheses, and their empirical consequences must be evaluated in a deductive system.

Despite the enormous difference between the analyses actually proposed for the natural languages phenomena at different points in time in the history of GG (the proposed computing systems), the general aim of chomskian linguistics has been remarkably consistent during all these years. We can say, in general terms, that Chomsky has obsessively pursued the same objective for 50 years, although from time to time he replaces the conceived theoretical device in order to attain the major task of his linguistics conception.

Deep down, what generative grammar intends to do is to construct a computing device, capable of forming and transforming representations, that can ‘simulate’ the linguistic knowledge a speaker of a natural language has ‘registered’ in his mind/brain.

It is this ‘nucleus’, constantly present in the 50 years of generative grammar history, that allows us to say we are before one and the same research program, in spite of the various deep changes the theoretical device (the computing system) has undergone.

2.2 The first proposal of generative grammar: LST and SS.

It is largely accepted that the GG history starts in 1957, with the publication of

⁵ The term ‘representation’ assumes many meanings in the linguistic literature. I am using it here to refer to formal objects of the theoretical construct that correspond to the things that belong to the modeled ‘reality’. In other words, the states of the mind/brain are, in our case, *represented* by expressions of a formal language.

Syntactic Structures (Chomsky 1957 – hereafter SS). Chomsky’s previous works – either because they circulate only among non-linguists, or because they had a limited circulation – exerted little influence in the development of the chomskian program. Chomsky’s M.A. dissertation (1951), for instance, was almost completely ignored by the linguistic community, although it has called some attention outside that area⁶. Other articles prior to that (Chomsky 1953 and 1955a) were more oriented to logicians and philosophers than to linguists: the former was published in the *Journal of Symbolic Logic*, a periodical rarely read by linguists; and the latter, although published in *Language*, supported a dispute with the Israeli logician Yehoshua Bar-Hillel about the applicability of symbolic logic developments to linguistic studies.

Around the mid-fifties, Chomsky finished writing an extremely pretentious book – *The Logical Structure of Linguistic Theory* (Chomsky 1955c – hereafter LSLT) – in which he put together - though as separate chapters - his Master’s dissertation and his PhD thesis (Chomsky 1955b) and cast the basis of a ‘new linguistics’. The book did not succeed in raising the editors’ interests and remained filed in microfilm form till its publication in 1975, as if it had just a historical value. In Chomsky’s words:

As for how LSLT was received, there is not much I can say about it. I have already mentioned that I was not aware I was doing linguistics. Therefore, the lack of interest or reaction from linguists was not surprising to me. I put LSLT forth to the MIT Press for consideration... but they refused it. They were right, I believe, because the contextual forces at the time were rather unfavorable for a general book on a subject like that. I also submitted a technical article on a limited part of the question to the Word magazine - following Roman Jakobson’s suggestion - but it was refused and sent back to me by mail. At the time I had little hope about having such work published, at least in a linguistics magazine, but quite frankly, I did not care much about that. (Chomsky 1977, p. 121)⁷

Then we come to 1957 and to *Syntactic Structures*. This book is a collection of notes from an undergraduate course Chomsky taught at the MIT and, according to Chomsky himself, does not reflect fairly his linguistic thoughts at that time. In his words:

You know what Syntactic Structures was. It was course notes for an undergraduated course at MIT. Van S showed up here once and took a look at some of my course notes for the undergraduate course I was teaching and said I ought to publish it. Since I had not been published anywhere, I said, why not, and that is what Syntactic Structures was. In fact, Syntactic Structures is a very misleading book. (Chomsky 1982a, p. 63)

Chomsky’s ideas only start to spread among linguists after the publication of *Syntactic Structures* (SS hereafter) and, mainly, after a long review that Robert Lees - a linguist who had a certain prestige in the community - published in *Language* (Lees 1957). So, the theory of this first period came to be known as ‘Syntactic Structure Theory’.

⁶ And this caused Chomsky to open his heart and say that his work called Bar-Hillel’s attention, for example, who said he was extremely interested in it. Also Quine were interested in the methodological aspect. But among linguists, no one showed any interest in that type of work. (Chomsky 1977, p. 120).

⁷ Translated from the Portuguese version by the author.

At this point, the nucleus of the chomskian program has already been fairly established, although not yet fully explained. The best way of capturing this nucleus is by investigating the main disagreements that Chomsky and some of his disciples understood to exist between their proposals and the proposals of the mainstream linguistics at the time: the American structuralism from a bloomfieldian lineage.

Chomsky's proposal, embodied in the SS theory, differ from structuralism in some important points:

First, the object of study of structuralism was *language*, understood as '*the totality of statements that can be uttered in a linguistic community*', according to Bloomfield (1926, p.47). It was the linguist's task to describe that language, and that would be done from the starting point of a 'representative corpus' of raw language data, which was described in detail with the tools offered by the 'discovery procedures'. What calls Chomsky's attention is the need to assume the existence of something prior to the structuralists' language: the capacity the speakers have to produce exactly the statements that *can* be produced. In other words, Chomsky moves the fundamental question of linguistic theory to determining the rules that govern those 'representative corpora', which lose their status as starting point of the linguistic theory and become its target. For Chomsky, the linguistic community has a shared knowledge about those utterances that can and those that cannot be produced, and it is exactly this knowledge that must be described and explained by the linguistic theory. The 'representative corpus' is the result of this knowledge and to take it as a starting point is methodologically uninteresting. According to Chomsky, a good indication of the existence of such knowledge lies in *linguistics creativity*: the speakers' ability of producing and understanding sentences to which they have not been exposed before.

We can then say that one could devise in Chomsky's initial works a psychological object for linguistic studies, although Chomsky did not put it explicitly like that. In other words, 'the state of mind/brain' mentioned above had already been glimpsed.

Another point of disagreement between American structuralism and the SS theory concerns the definition of the *aims* of those theories. While the structuralist theories were, generally speaking, explicitly *descriptive*, the SS theory intended to be *explanatory*, in the sense that the phenomenon should be deduced from a set of general principles⁸. The adoption by Chomsky of a hypothetical-deductive model of science presents deep implications to his program's procedures. It is no longer a question – as in structuralism – of *describing* data that reveal themselves to the linguists' perception, but a question of finding the *general principles* from which the descriptions of observable data can be logically derived. With Chomsky, the theoretical aspect takes precedence over the empirical aspect. It is not surprising, however, that a large amount of Chomsky's work, at that time, discusses formal languages and tries to define the formal nature of natural languages as compared to the languages used by logicians and mathematicians⁹. It is the program's heuristics determining the priorities and, clearly enough, Chomsky realizes that what is fundamental at that moment is the definition of a

⁸ Chomsky explicitly assumes a deductive-monological perspective in his characterization of what might be an explanatory theory (see, for example, Chomsky 1997, p.106).

⁹ It is in this period that Chomsky develops a classification and a typology of languages that is still used by logicians, mathematicians and computer scientists – the so-called Chomsky's Hierarchy. With his hierarchy, Chomsky intends to show that natural languages present properties that cannot be represented by grammars that would be perfectly adequate to account for the languages used by logicians. For example, while logicians' languages can be represented by formal systems (free-context grammars) that construct expressions without taking the context into account (the adjacent expressions), natural languages allow to construct expressions that can only be represented by more powerful formal devices – the transformational grammars.

formal notion of grammar, understood as a computing system (a *generative grammar*, in the technical sense of the word), which, like a kuhnian *paradigmatic sample*, could be used as a tool for describing natural languages phenomena. In other words, to effect a proposal of describing the speakers' implicit knowledge and do so in the frame of an explanatory theory, Chomsky needs to construct a formal device (a generative grammar) capable of accounting for the well-formation rules of any language L and of relating this formal device to some set of general principles (that may determine what can stand for a 'generative grammar' of languages in general).

Then a division emerges classifying the tasks in linguistics into two groups: the construction of grammar for particular languages, and the construction of general principles for language competence ('universals'). The process of constructing particular grammars requires formal devices that are powerful enough to account for details and specificities of languages: the process of construction of general principles must ignore the specificities and search for the 'universals'. To a certain extent, the tension between these two processes will be responsible for the changes in the models of analysis that will recur in time during the last 50 years.

The first model of analysis proposed by Chomsky is presented at length in LSLT and SS and consists basically of a sophisticated form of immediate constituent grammar plus a transformational component. This model of analysis presents two main components: one *forming* expressions, which is a generative version of the immediate constituent grammar, developed and presented by Rulon Wells (1947), among others; and another which *transforms* expressions, and that, at least in part, stands on the notion of 'transformation' developed by Zellig Harris, who was Chomsky's tutor for his doctorate degree¹⁰. Besides these basically syntactic components, there is a morphophonemic component, which attributes phonological readings to the output of the transformational component.

The basic notion is that of *linguistic level*.

A language is an enormously complex system. Linguistic theory attempts to reduce this immense complexity to manageable proportions by the construction of a system of linguistic levels, each of which makes a certain descriptive apparatus available for the characterization of linguistic structure. A grammar reconstructs the total complexity of a language stepwise, separating out the contribution of each linguistic level.

(Chomsky 1955c, p. 63)

According to Chomsky in LSLT, a linguistic level is an L-system in which unidimensional representations of utterances are constructed. Each level presents a fixed and finite 'alphabet' of primitive elements. Through a *concatenation* operation, we can obtain sequences of elements that will be called *chains in L*. In the process of linguistic analysis, a set of chains called L-markers is constructed at each L-level, which will be attributed to the sentences of the language under analysis. The L-marker of a given sentence S must contain all the structural information referring to S at L-level. The relationship between various levels L_1, L_2, \dots, L_n is established through mapping operations, which associate the elements on a level to the elements on other levels. If we organize the various levels into a hierarchy, we can think of a sequence of mappings (from L_1 to L_2 , from L_2 to L_3 , and so on so forth up to L_n) till we get to a last level, which associates L-markers to the sentences of the language.

¹⁰ See, for example, Harris 1957.

We will find it necessary to distinguish at least the following levels for linguistic description: phonemes (Pm), morphemes (M), words (W), syntactic categories (C), phrase structure (P), and transformations (T). (Chomsky 1955c, p. 66)

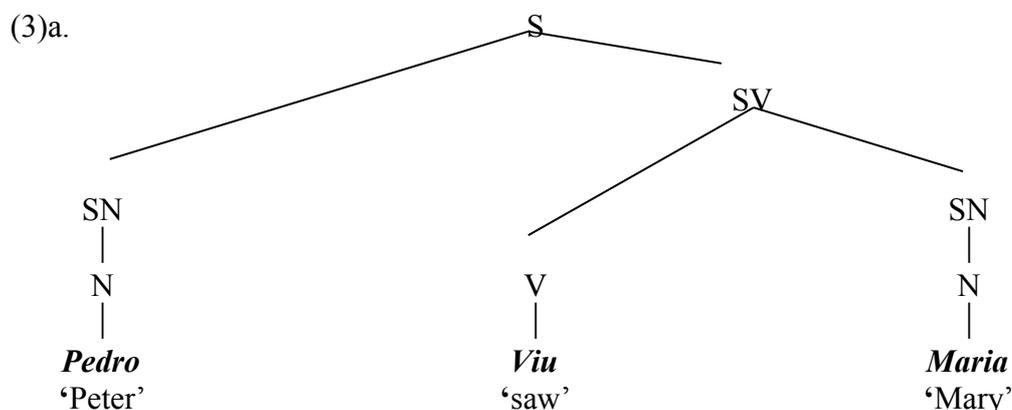
It is then within the scope of grammar to attribute at least six representations – one for each level – to each sentence of the language. In other words, the grammar must assign to each sentence a formal representation (an ‘expression’ of language in which the computing system is written), on each of the linguistic levels, which may represent (simulate, model) the properties of the sentence referring to that given level.

A grammar of a language must tell us exactly what are the grammatical sentence tokens, and exactly how these are represented on each level. (Chomsky 1955c, p. 99)

Let us now look at the form a grammar must take in order to carry out these tasks. First, it is necessary that the grammar establishes for each sentence-token a sequence of representations $\langle R_1, \dots, R_n \rangle$, where R_1 is a representation of Sentence, R_n is a phonetic representation and R_2, \dots, R_{n-1} are intermediate representations.

We can generate these representation sequences by rules of the form (1) $X \rightarrow Y$ interpreted as the instruction “rewrite X as Y.” We call each such rule a conversion. (...) We say that the sequence $\langle R_1, \dots, R_n \rangle$ is a derivation of R_n , generated by a set C of conversions, if R_1 is Sentence and for each i ($1 \leq i \leq n$), R_{i+1} follows from R_i by one of the conversions of C. (Chomsky 1955c, p. 114)

The basic idea is that a sentence such as *Pedro viu Maria* ‘Peter saw Mary’, like any other sentence in the Portuguese language, is assigned a set of representations and that these representations may be constructed in a sequence, one after the immediately previous one. Let us suppose that the system starts saying that *Pedro viu Maria* is a sentence. From that point, the system should say how the sentence consists of syntactic categories, showing its structure. Our sentence, for example, consists of a noun phrase (NP) followed by a verb phrase (VP); the noun phrase on its turn consists of a verb (V) followed by another NP. The NP’s consist of names (N). Thus, at phrase structure level, the sentence *Pedro viu Maria* will be assigned a representation that may have one of the following two forms (which are absolutely equivalent):



(3)b. (((*Pedro*)_N)_{SN} ((*viu*)_V ((*Maria*)_N)_{SN})_{SV})_S
 ‘Peter’ ‘saw’ ‘Mary’

Each of the words *Pedro*, *viu*, and *Maria* will be assigned a representation at the morphological and phonological levels, which will allow to obtain a morphological representation for the sentence. And so on so forth, till we end up with a fully represented sentence in all six linguistic levels.

Each conversion must be marked somehow in order to indicate which linguistic level it belongs to, and this can be done by grouping the conversions of the same level and by establishing a certain order of application.

Thus we come to the well known grammar of Syntactic Structures:

E: Sentence

F: $X_1 \rightarrow Y_1$
 · ·
 · ·
 · ·
 $X_n \rightarrow Y_n$ } Phrase-structure

T_1 }
 ·
 ·
 ·
 T_j } Transformations

$Z_1 \rightarrow W_1$ }
 · ·
 · ·
 · ·
 $Z_m \rightarrow W_m$ } Morphophonemics

To produce a sentence from such a grammar we construct an extended derivation beginning with Sentence. Running through the rules of F we construct a terminal string that will be a sequence of morphemes, though not necessarily in the correct order. We then run through the sequence of transformations T_1, \dots, T_j , applying each obligatory one and perhaps certain optional ones. These transformations may rearrange strings or may add or delete morphemes. As a result they yield a string of words. We then run through the morphophonemics rules, thereby converting this string of words into a string of phonemes.

(Chomsky 1957, p. 46)

This, in short, is the model of analysis Chomsky proposed in the late fifties.

The hypothesis put forward then is that this linguistic analysis device would be an adequate representation of the computing system, present in the speakers' mind/brain, able to determine, at least partially, its linguistic components. That is, assuming the affirmations of the nucleus, the heuristics determined the construction of a representation hypothesis of the computing system, which, until disproved, was

considered *the grammar present in the speakers' main/brain*.

It is easy to realize that this model is not far away from the models resulting from the American structuralism. Chomsky himself does not see any fundamental differences between his model and Bloomfield's and Harris's distributional model: at the end of chapter IV in LSLT, after presenting his model of grammar, Chomsky states:

We will refer to linguistic analysis carried out in these terms as "distributional analysis." This usage seems to me to correspond to the practice of what has been called distributional analysis.
(Chomsky 1955c, p. 127)

In what concerns the notion of linguistic level, the chomskian approach is not markedly different from the structuralist approaches: Chomsky refers to Hockett (1955) 'for a similar approach to linguistic levels' (1955c, p.97). The only point in which Chomsky's proposal seems to fall far apart from the structuralist approaches is in what concerns the number of levels necessary for a linguistic description.

Our main conclusion will be that familiar linguistic theory has only a limited adequacy – i.e., that it is attempting to do too much with too little theoretical equipment. (...) We will argue that the remedy for these deficiencies is not to be found in the extension of the distributional basis for linguistic theory to include meaning, situational context, etc., nor, apparently in the introduction of probabilistic and statistical conceptions. Instead, a new level of transformational analysis is proposed as a higher level of linguistic structure. It will be shown that the theory of transformational analysis can be formulated in the same completely distributional terms that are required anyway for lower levels, and that a large and important class of problems that arise in the rigorous application of familiar linguistic theory disappears when it is extended to include transformational analysis.
(Chomsky 1955c, p. 64)

Apparently, according to Chomsky, the addition of a transformational level to the 'known linguistic theory' (in other words, American structuralism) is enough to turn it adequate. Lees, by the way, assume the same position, with an exaggerated optimism, He declares:

*Chomsky...has been led to set up a whole level of grammatical transformations to deal with **all the difficulties** encountered in trying to state explicitly a complete and simple immediate-constituent grammar.*
(Lees 1957, p. 52 – bold type added by the author)

However, it is somehow strange to consider that Chomsky constructs a set of transformational rules aiming at overcoming the deficiencies of the AS models since in Harris (1952) transformational rules were already used for the analysis of natural languages, and Harris is one of the most characteristic representatives of AS. We must make it clear then how Chomsky's notion of transformation differs from Harris's transformation.

Harris's notion of transformation lies in the notion of *sentence form*. Harris arrives at this notion by the definition of variables having word classes as domains (variable N has as its domain the class of name, for example). From there on, it is

possible to define sequences of well-formed variables¹¹. Harris calls these sequences of well-formed variables *sentence forms*.

It happens, however, that not all sentences obtained by attributing values to the variables of a sentence form are equally acceptable. For example, the sentence *o homem pensa* ‘men think’ is more acceptable than *a pedra pensa* ‘stones think’, although both belong to the same sentence form. Each sentence form, then, has a gradation of acceptability for the values of its variables and such gradation in fact characterizes the sentence form.

Harris defines, then, *pairs* of sentence forms, made up exactly of the same variables, differing only by some univocally determinable characteristics (difference of order between variables; constant presence or absence of a certain element; etc.). If two sentence forms belong to a pair thus defined and, in addition, present the same acceptability gradation, Harris will say that they are in *transformation relation*. For example, the sentence forms N_1+V+N_2 and $N_2+ser+V-do\ por+N_1$ constitute a pair of forms that (i) present the same variables; (ii) present constant elements in one of them (*ser, -do, por* ‘be, -ed, by’), and (iii) present a change in order between N_1 and N_2 . In as much the acceptability gradation of both sentence forms is the same, Harris will say that these two forms are in a transformation relation named *active/passive*.

According to Harris, then, a transformation is nothing more than a class of pairs of sentences. As Milner (1973, p. 191) points out, Harris’s transformations are relations that can be expressed in a class language and can be labeled *relations-in-extension*.

Let us now see what transformations consist of in Chomsky. He starts from the definition of ‘analyzable’ predicate. Let us imagine a sentence consisting of a sequence of elements t ; let us imagine that Q is a syntagmatic indicator (a ‘tree’, as the one in (3b)) representing structure t ; let us now suppose that t can be subdivided into successive segments t_1, \dots, t_n so that each t_i is linked to a node A_i in Q . Given these conditions, we can say that t is analyzable in $\langle t_1, \dots, t_n; A_1, \dots, A_n \rangle$ from the point of view of Q .

Let us take the sentence *Pedro viu Maria* ‘Peter saw Mary’, for example. We can represent its structure by means of the syntagmatic indicator in (3b). The sentence is divided into three successive segments (*Pedro, viu, and Maria*) and for each segment we can find in the syntagmatic indicator a node to which it is linked: *Pedro* is linked to N, *viu* is linked to V, and *Maria* is linked to N. We can then say that the sentence *Pedro viu Maria* is analyzable into $\langle Pedro, viu, Maria; N, V, N \rangle$ ¹².

According to Chomsky, a transformation has a *domain* (or a ‘structural condition’) and an *effect* (or ‘structural change’). The domain indicates the class of linguistic expressions that can undergo transformation and is specified by a sequence of symbols $\langle A_1, \dots, A_n \rangle$, which are symbols of nodes of a syntagmatic indicator. In order for a linguistic expression to be in the domain of transformation it is necessary that it be analyzable into $\langle t_1, \dots, t_n; A_1, \dots, A_n \rangle$. The transformation effect, on its turn, is described by a rule that projects the starting sentence (the set t_1, \dots, t_n of the sentence that will undergo transformation) into the target sentence (i.e., the already transformed sentence). Let us take a *passive* transformation for example. Its form in Portuguese would be roughly the following¹³

¹¹ In Portuguese, for instance, the sequence N+V+N is well formed, whereas the sequence V+N+V is not.

¹² The sentence *Pedro viu Maria* is analyzable, from the point of view of the same syntagmatic indicator, into $\langle Pedro, viu Maria; N, SV \rangle$, provided we consider the SV internal structure.

¹³ This is just an outline of the passive transformation in Portuguese. It is necessary to underline that this type of transformation was already abandoned in the initial revisions of the GG and, as much as I know, it has never been

Domain: $N_1 \text{ Aux } V \text{ } N_2 \quad \Rightarrow$
 Effect: $N_2 \text{ Aux} + \text{ser } V + \text{-do } \textit{por} \text{ } N_1$
 ‘be’ ‘-ed’ ‘by’

The domain tells us that expressions can undergo transformation (the sentence *Pedro viu Maria*, for example, is within the transformation domain¹⁴); the effect tells us which are the changes the starting sentence (let us say, *Pedro viu Maria*) must undergo in order that the target sentence can be obtained. Basically, the changes consist of the permutation of subject (N_1) and direct object (N_2); the addition of verb *ser* ‘to be’, as ‘tense bearer’ auxiliary; the addition of the past participle ending to V ; and the addition of the preposition *por* ‘by’ before N_1 . After applying the rule to the sentence *Pedro viu Maria*, we will obtain the target sentence *Maria foi vista por Pedro* ‘Mary was seen by Peter’.

According to Milner, what Chomsky defines with his transformation notion...

c'est une entité spécifique, la règle de transformation et non une classe de paires de phrases; le fait que deux phrases soient en relation est envisagé comme une propriété de la paire, distincte de la paire elle-même, et dont la règle prise dans son ensemble est le symbole.

(Milner 1973, p. 104-105)

For Milner, chomskian transformations cannot be dealt with in a class language and characterize *relations-in-intention*.

It is easy to see that there are many similarities between the two concepts of transformation. According to Milner, what make them different are the approaches: *extensional* in Harris and *intentional* in Chomsky.

In extensional terms, a transformation relation is totally determined by the pair of sentences it relates and the task of the grammar is simply to establish those pairs. Harris states quite clearly that, basically, transformational analysis is not a means of determining the structure of each sentence considered separately, but rather a way of grouping the sets of sentences into pairs (Cf. Harris 1968, p. 68).

From an intentional perspective, on the other hand, given a pair of sentences, we can imagine various relations (rules) between them, specified differently by the *analyzable* predicate. The scientist's task is much more complex than it would be from an extensional perspective.

We could always argue that the extensional and intentional theories are equivalent, since it seems possible to establish corresponding tables that relate *properties* and *classes*, *relations-in-extension* and *relations-in-intention*, etc. For example, the statement ‘sequence *t* is analyzable into...’ is nothing more than the intentional version of the statement ‘sequence *t* is a member of the sentence form...’. However,

Il exist bien des cas où les paires sont parfaitement connues, mais où la règle de transformation intensionnelle ne peut être formulée parce que le prédicat « analysable » ne peut être précisé à coup sûr.

Les exemples réels abondent et certains d'entre eux sont très connus : ainsi les paires phrases actives/phrases passives sont parmi les mieux

studied seriously in relation to Portuguese data.

¹⁴ The node Aux (auxiliary) serves, basically, to ‘transport’ verb tense. In the case of *Pedro viu Maria* the Aux is [+pass].

attestées de la tradition grammaticale ; en termes extensionnelles, la transformation passive est donc une de plus certaines et des mieux décrites. En revanche, si on envisage les données du point de vue du prédicat « analysable », de graves obscurités apparaissent : quelle est la nature exacte du complément d'agent, quelle est la condition exacte qui spécifie le domaine de la transformation (la nature transitive du verb est-elle suffisante ? D'autres facteurs interviennent-ils : par exemple, un Adverbe de Manière fictif ? cf. Aspects, p. 145-150), etc. Bref, la règle de transformation passive est une des plus mal connues de la grammaire intensionnelle.
(Milner 1973, pp. 198-199)

We cannot then speak about equivalence between the two perspectives because the relationship between them is asymmetrical: given a well-formed intentional theory, we can easily obtain an extensional counterpart, but the opposite is not true.

There is still another difference between the two notions of transformation that is fundamental for its comprehension (and that is ignored by Milner). It is the fact that not everything Chomsky lists as transformations are sentences. The harrisian notion of transformation pairs sets of sentences (sentence forms); the chomskian notion of transformation maps a structure of a sentence onto a given derivational stage in another structure of the *same sentence* obtaining another derivational stage. Chomskian transformations definitively are not devices constructed for obtaining pairs of sentences although, through them, we can justify the pairs of sentences our intuition recognizes in the language.

It is then clear that Chomsky really innovates when he proposes his intentional theory of transformations. The notion of transformation – and the role this notion plays inside the model – in fact opposes Chomsky to AS. It thus justifies the emphasis Chomsky and those who publicize his theories (such as Lees) give to this notion.

In spite of the innovation represented by the introduction of the 'transformational level' into the linguistic analysis, generally speaking, the chomskian descriptive model does not deviate much from the structuralist tradition.

Thus, we find GG in its initial years having to face an inconsistency that, although not perceived as such at that time, will require deep changes in the form of the theory: the program's proposal is different from that of the structuralist program, but the available analytical tools are basically the same. That is, there is a mismatch between what is intended to do and what, in fact, ends up being done.

So, it is not surprising to see that alterations are soon included in the descriptive model by assuming new auxiliary theories.

2.3. Generative Grammar in the sixties: the standard theory.

The first ten years of generative grammar were its 'heroic years' in which the combat with the forces of American structuralism dominated the scene¹⁵. The SS theory proceeds having its presuppositions explicitated and some of its theoretical devices changed or replaced in order that it could manage - in better conditions - to carry out the task the theory imposed, namely the description of a computing system able to define,

¹⁵ Despite being very similar, as we saw, GG and the American structuralism soon start out a dispute for space and prestige inside the North-American academic institutions. As this dispute is a matter of sociology of science, rather than of philosophy, and the present work intends to have a basically philosophical character, I am not making any attempt to explore here this bias of the history of science.

generatively, natural languages. In the mid sixties one could already consider that the battle had been won: the generative program was clearly dominating among the North-American linguistics and began its expansion to other places, by conquering new followers outside the USA.

Aspects of the Theory of Syntax (Chomsky 1965) is a great synthesis of the changes that were made in the SS theory. Besides the first great clarification of the program's postulates - presented in the 'Methodological Preliminaries'¹⁶ - Chomsky develops a new configuration for his grammar: a model that came to be known as the 'Standard Theory'.

This new theory changes significantly the descriptive model and clarifies a series of postulates that, in the previous theory, had been presupposed in a rather obscure way or, though resulting from those presuppositions, were not presented clearly. It is during this period, for example, that the question of innatism is raised as a working hypothesis, with the resulting strong psychologization of the grammar¹⁷. It is also in the same period that powerful auxiliary theories were developed, which will allow for a better descriptive and/or explanatory performance of the program: the lexicon becomes relevant and receives its first consistent theoretical formulation; the 'deep structure' appears, having as its strongest consequence the outset of a deeper concern about semantics (concern that provided for the emergence of a series of more or less heterodox alternative theoretical formulations)¹⁸.

It is also not surprising that a high number of discussions focused the *transformational component* of the model of analysis. Given a set of linguistic phenomena, if what differentiate a generative analysis from the one carried out by a structuralist is the presence of transformations, it is obvious that the heuristics should determine a more careful exam of that notion aiming at the solution of formal problems that might emerge there.

In the standard theory, the form of grammar - which in a way represents the image that one had at that time about the functioning of the linguistic knowledge the speakers have registered in their mind/brain - is organized into three major components: a *syntactic* component, which is generative, in as much it is the only component that constructs representations, and two interpretive components; the *semantic* component; and the *phonological* component¹⁹.

The process of generating sentence starts from the syntactic component, which has the following basic structure: a base subcomponent (or simply BASE), which is the responsible for generating the *deep structures* (DS); and a transformational subcomponent, which converts the DS's into *surface structures* (SS). The base subcomponent contains (i) a set of rewriting rules (sometimes called *categorical component*), which, applied to the initial axiom S, generates tree structures 'labeled' with the symbols of the categories whose terminal nodes are not filled in; and (ii) a lexicon, which inserts lexical items into the terminal nodes of the tree. The base input is the axiom S and the deep structures are the output. The transformational component is assigned deep structures, as input, and, through transformational rules, converts them into surface structures. In a diagrammatical form, we would have:

¹⁶ First chapter in Chomsky 1965.

¹⁷ Linguists spent a lot of time discussing the Language Acquisition Device (LAD).

¹⁸ As Generative Semantics, for example (see Lakoff 1971, Kato 1974, Dascal 1978, and Galmiche 1979, among others).

¹⁹ The difference between the *generative* component and an *interpretive* component lies in the property that the generative component has of creating new representations, whereas the interpretive components only associate (pair, relate) representations among themselves.

The syntactic component specifies an infinite set of abstracts formal objects, each of which incorporates all information relevant to a single interpretation of a particular sentence.

(Chomsky 1965, p.16)

induced the generativists to search for syntactic solutions to semantic problems and to look for DS's that could represent *all* aspects of meaning in the sentences under analysis. This procedure led to the postulation of DS's that were more and more abstract and closer and closer to the semantic representations²⁰. The commitment with the hypothesis that the semantic interpretation occurs at DS level leads many linguists to conclude that everything one can consider as being part of the meaning of the sentence must be included in the DS. Thus, for example, all ambiguities observed in the sentences should be solved by postulating different DS's, not mentioning the phonological phenomena with semantic consequences, as focalization, presuppositions, and performatives, for example.

A series of analyses and a series of empirical and theoretical arguments were raised by the 'abstractionists', all leading to the same conclusion: it is necessary to postulate more abstract DS's that may represent more directly the semantic representations present in the sentences.

It is important to point out that the 'abstractionists' remained strictly inside the standard theory and, consequently, inside the GG program, what can explain the almost general acceptance of their analyses by the generative community, at least for a while.

With the expansion of the abstractionist posture, one comes to an almost complete decharacterization of the DS notion - as Chomsky imagined it - and it became meaningless to try and differentiate them from the semantic representations.

The 'abstractionists' tried – in an ordered way – to gather their ideas in what was considered at the time a new 'paradigm', which was labeled *Generative Semantics*²¹.

Chomsky's reaction did not take long to manifest and, in 1967, he proposed some changes in the standard theory in order to avoid the uncontrolled abstraction. Chomsky's reaction had some consequences. First, the breaking up with the 'abstractionists' is inevitable and a first group of linguists 'generated' inside the GG program emerges and they plunge into the task of constructing new research programs. Second, a new model of linguistic analysis is established. This new model came to be known as the *Extended Standard Theory* (EST)²².

As the name reveals, the EST is not understood as a new theory. It is just the old standard theory that received some new auxiliary theories; the most important of them is the *X-bar Theory*. The role of the lexicon is also changed; the lexical items, for example, start to be considered a bundle of traces, and many phenomena that were approached via transformations began to be approached via lexical relations. In all, however, the theory remains the same.

2.4 From rules to principles.

After the heat of the famous sequence of clashes with the Generative

²⁰ That is, the formulae (expressions of the formal system language) which were supposed to represent adequately the meanings of natural language expressions.

²¹ As I cannot present a more detailed analysis of this period within the scope of the present work, I suggest the interested reader to refer specially to Newmeyer 1980, Harris 1993 (and Borges Neto 1991) for details.

²² EST is initially proposed by Chomsky 1967 and developed in Chomsky 1968 and 1971.

Semantics cooled down, Chomsky and his associates could go back and think about the development of the GG program. The greatest legacy of Generative Semantics was the proliferation of theoretical devices and of the types of available rules. The immediate task the chomskians had to face then was that of restricting the descriptive power of the grammar in order to increase its explanatory power.

A certain ‘tension’ is felt in the GG program between the requirement of *descriptive adequacy* and the requirement of *explanatory adequacy*. Chomsky says that a theory is *explanatorily adequate* when it manages to successfully select a descriptively adequate grammar from a set of possible grammars (defined by the general theory) and from the *primary linguistic data*. In other words, the theory is explanatorily adequate when it reproduces the behavior of children acquiring language: from the raw linguistic data, they select a grammar among the possible grammars admitted by the innate component of their linguistic competence²³. Thus, in order to attain the explanatory adequacy, the available theoretical devices in the general theory (the grammar innate component theory) must be restricted, so that few grammars can be obtained and one is able to understand how a child quickly selects an adequate grammar for the data available to him/her.

On the other hand, in order to attain the descriptive adequacy, that is, in order to construct grammars to all natural languages, the available theoretical devices must be rich and varied enough to cover all the richness and diversity of natural languages.

The conflict between these two requisites of adequacy is obvious and the search must go for a theory that is, at the same time, rich enough to account for the variety of languages and restrictive enough to provide a small number of possible grammars.

In the mid sixties it is already possible to find restrictive proposals to the descriptive power of transformational rules. John Robert Ross’s PhD Dissertation (Ross 1967), postulating the ‘islands’ (syntactic configurations that prevented the extraction of elements), is a good example of this concern. But it is with Peters and Ritchie’s works, in the early seventies, in the heat of the debates between the GG and Generative Semantics, that the need to restrict the power of those rules becomes urgent.

Peters and Ritchie (1969, 1971, 1973) demonstrate that the weak generative capacity of a grammar that includes transformational rules, as those proposed at that time, is equivalent to an unrestricted rewriting system (a Turing machine). This means that a transformational grammar of any natural languages do not reveal anything about the structural characteristics of that language, but only states that that language - understood as a set of sentences - is recursively enumerable (and, consequently, can be generated by calculus). Peters and Ritchie’s works show that the major problem of transformational grammars was not the proliferation of rules or categories but the lack of strong restrictions on the functioning of those rules. Without restrictions, the transformations – that could eliminate, create, exchange, move, or change elements – were useless as exposing devices of linguistic structures. Applying the adequate deletions, movements, and additions, one could start from any sentence and arrive to any other sentence.

The task of proposing restrictive conditions on the functioning of the rules becomes the priority of the program. It seems clear that here we have a case of *creative change* in the heuristics. The focus is not placed on the descriptive adequacy any longer

²³ The innate component – Language *Competence* – is a set of general principles about the nature of the linguistic representations, principles that can be manifested in a whole range of alternative realizations (the parameters). According to Chomsky, children apply these ‘parameterized’ principles to construct the grammar of their language, provided they are exposed to the data.

but on the explanatory adequacy. Instead of having a proposition of computing systems as the most important task, the restrictions of the systems previously proposed become the priority. Obviously, the orientation of the theory - its 'plan of development'-undergoes a substantial change.

It is possible to see two main tendencies in this process of restricting the power of rules. The first tendency is to impose general conditions to the application of rules: Ross's Island Conditions (1967), Emonds's structure-preservation restriction (1979), and Chomsky's conditions (condition over specified subject, condition over sentence with tense and subjacency condition) are all conditions over the *application* of transformational rules. In other words, they tell us that transformations only apply if certain conditions are present. For example, Emonds's structure-preserving restriction tells us, among other things, that one cannot delete anything in the structure that cannot be structurally recovered: it is only possible to delete the subject of the sentence in Portuguese because the verbal inflection allows to recover structurally the subject position; in a language without verb inflection, such as English, subject deletion is hindered.

Another tendency, besides restricting the application of rules, proposes a strong restriction onto the *number* of available rules. This tendency emerges with Chomsky (1976a) and becomes a dominating tendency in GG from then on. In Lobato's words:

*Restriction to the number of transformations has been a characteristic of the chomskian generative theory since 1976 ('Conditions on rules of grammar'), when the transformational component was said to consist of just **two** rules*

NP- move,

Wh- move.

(Lobato 1986, p.337)²⁴

This second tendency is interesting and deserves some attention.

When Chomsky proposes that the NP postposition and NP anteposition transformational rules be replaced by only one rule of *NP-move*, on the one hand, he manages to reduce the number of available rules, but, on the other hand, he obtains a rule that is so general that any NP is possible to be moved from anywhere to anywhere else. How is it then possible to maintain the reduction in the number of rules without losing sight of the general need for restricting the grammar as a whole? Chomsky's solution is to propose an interaction of rules with a set of general principles about grammar. The solution proves to be so operational that Chomsky goes ahead and reduces even more the transformational component, maintaining only one rule: the MOVE ALPHA rule.

As the transformational component is reduced to only one rule, and considering that this rule is optional, in order to permit the generation of alternative surface forms for one and the same underlying structure, it is necessary to find out devices that not only hinder undesirable moves but also devices that force the move in cases in which it should be obligatory.

Chomsky's solution lies in the establishment of some new auxiliary theories: the Case theory, the Thematic Role theory (Theta theory), the trace theory, the empty category theory and the Binding theory besides the X-bar theory, already on. All these theories impose conditions onto the possible representations and, consequently, both force and hinder moves.

²⁴ Translated from the Portuguese original by the author.

I am not going into long analyses within the scope of the present work. But I would like to show just how these auxiliary theories work.

The trace theory, for example, developed in Chomsky (1976a), postulates that all moved element leaves a *trace* in the place it left, and this trace acts as a full element for the purposes of syntactic rules. For example, the analysis of sentence (4) proposes it be derived by transformation from structure (5) through the move of the clitic *nos* to a position next to the verb of the main clause²⁵.

- (4) *Paulo nos viu examinar a garota*
 ‘Paul saw **us** examine the girl’
 (5) *Paulo viu [nos examinar a garota]*
 ‘Paul saw [**us** examine the girl]’

It is not possible, however, to obtain (6) from (7) because the clitic *a* would have to ‘fly over’ the subject of the subordinate and this is not allowed by the *Conditions over Specified Subject*²⁶.

- (6) * *Paulo a viu nós examinar*
 ‘Paul saw us examine **her**’
 (7) *Paulo viu [nós examinar a]*
 ‘Paul saw [us examine **her**]’

Let us consider a fairly more complex case. Sentence (8) would be obtained from the structure present in (9) by the move of the subordinate subject to the position of subject of the main sentence and by the anteposition move of the clitic *os*.

- (8) *Paulo parece os ter examinado*
 ‘Paul seems to have examined **them**’
 (9) Δ *parece [Paulo ter examinado os]*
 ‘ Δ seems [Paul to have examined **them**]’

Sentence (10), however, which should behave in the same way, is not grammatical, although there is apparently no other reasons for the application of the specified subject condition.

- (10) * *Paulo os parece ter examinado.*
 ‘*Paul seems to have examined **them**’

The solution to the problem brought by the trace theory suggests that between sentence (9) and sentence (10), we should have an intermediate structure (11) that, having its trace in subject position, hinders the move of the clitic outside the subordinate sentence.

- (11) *Paulo parece [t ter examinado os]*
 ‘Paul seems [**t** to have examined **them**]’

Clitic moves are allowed inside the subordinate sentence, and so, from structure (9) we could obtain structures (12) and (13) with no difficulties.

- (12) *Paulo parece [t tê-los examinado]*
 ‘Paul seems [**t** to have them examined]’

²⁵ These analyses are borrowed from Quicoli 1976.

²⁶ Which is one of the conditions on transformations proposed by Chomsky.

- (13) *Paulo parece [t os ter examinado]*
 ‘Paul seems [t to have examined them]’

The trace hypothesis hinders the undesirable move maintaining the generality of the move rule.

The Case theory, on its turn, determines that all full NP (i.e., a morphologically realized NP) receives a CASE²⁷; and, in addition, establishes the contexts²⁸ in the structure in which an NP can receive a CASE. Thus, an NP generated in a position that does not allow Case attribution must necessarily be moved to a position in which Case may be assigned. For example, in (14) the subject position of the main sentence is empty in the underlying sentence and, in order that a well-formed sentence can be obtained, it is necessary that the subject of the subordinate occupies the subject position in the main sentence, producing (15).

- (14) *Δ parece [João estar alegre]*
 ‘Δ seems [John be happy]’
 (15) *João parece [t estar alegre]*
 ‘John seems [t be happy]’

Now, that move is not obligatory in all cases, since we could obtain (16) from (17), for example.

- (16) *CV parece [que João está alegre]*
 ‘CV seems [that John is happy]’
 (17) *Δ parece [João está alegre]*
 ‘Δ seems [John is happy]’

Instead of proposing, in an *ad hoc* way, that certain moves are obligatory, it is preferable to establish a general principle, such as the attribution of Case to full NP’s - like *João* - and determine the contexts in which these attributions will take place. We can see in our examples that in (17) the NP *João* is assigned Case because it is the subject of a sentence with a finite verb (the subordinate sentence) whereas in (14), since the verb in the subordinate sentence is an infinitive verb, the NP *João* is not assigned Case, requiring, however, to be moved to the subject position in the main sentence in order to be assigned the nominative Case (the main verb *parece* is a finite verb and, consequently, allows the attribution of Case). Notice that the move remains optional. However, the absence of move, in (14), turns the structure agrammatical.

Summing up, one of the theories – the trace theory – restricts the number of places to which the moved element may go to, whereas the other – the Case theory – turns certain moves obligatory.

Chomsky (1976a and 1976b) notices that these conditions produce a more general effect than simply that of regulating moves: the conditions also hinder certain relationships between elements in cases in which there was no move. So, the proposal is that the conditions work also in the regulation of the interpretive relations. This way, it is possible to show a structural parallelism between the relationship between the trace and its antecedent (consequence of the move rule) and the relationship between certain anaphoric pronouns (reflexive, reciprocal and PRO) and its antecedents (consequence of interpretive rules).

²⁷ In some of the latest versions of GG, it is also possible to attribute cases to NP’s not realized phonetically.

²⁸ The attribution of Case happens as follows: (i) the sentence with a finite verb attributes a nominative case to its subject; (ii) the verb attributes an objective case to its complement; and (iii) the preposition attributes an objective case to its complement.

What can be perceived hereon is that it becomes impossible to admit that the move rules are free, that is, that they are not submitted to conditions. The conditions on moves cease to exist, and all their effects turn to be obtained through *interpretive conditions*.

The main consequence of this is the substitution of what we can call *derivational perspective* by a *representational perspective*. We are facing a new creative change in the program's heuristics.

From a derivational perspective, the various representations of the linguistic levels (they may be 'phonemes (*Pm*), morphemes (*M*), words (*W*), syntactic categories (*C*), phrase structures (*P*), and transformations (*T*)', as Chomsky intended in LSLT; they may be 'Deep structure, surface structure, phonetic form and semantic representation', as in the standard theory) are *derived* from one another through rules. The grammar is strictly directional, that is, the various levels of linguistic analysis are approached – and receive representations – in a given order. Thus, until the mid seventies, all proposals of grammars made inside the GG were derivational.

From the *representational* perspective, on the other hand, the various representations are not related by derivation: they are just representations of structural properties resulting from the theories restricting grammar. P-structures, for example, can turn to be understood as a 'pure' representation of the grammatical functions relevant to the attribution of thematic roles and, in this sense, as an 'abstraction' of the S-structures. Grammar is not directional. As Lobato states:

This change of approach leads to a switch in the interpretation of what may be 'generated by the base'. In the previous versions of the theory, this expression meant 'derived from S by successive applications of syntagmatic rules e by using the lexical substitution rule'. Now it means 'be projected from the lexicon, from X, according to the UG [Universal Grammar] principles and the parameters established by a certain language'. This new perspective allows then to consider an S-structure generated by the base, and the move-ALPHA is a property of the S-Structures and not, from this point of view, a rule that converts P-structures into S-structures (cf. Chomsky 1982b: 33). Likewise, any other level of representation can be seen as 'derived by the base', since any level of representation is determined by the establishment of the UG parameters (Chomsky 1982b: 14).

(Lobato 1986, pp. 403-404)²⁹

Having completed the cycle of substitution of rules by principles, Chomsky finds himself involved by a grammar theory which is different enough from the previous ones to justify a new label: *Principles and Parameters Theory*.

I am not going into details about the Principles and Parameters Theory (also called, for some time, *Government and Binding Theory*). It suffices to say that this is the theory used at this beginning of the XXIst century for the study of the syntax of natural languages. However, it is worth mentioning some movements that have been observed inside the GG program – especially the so-called Minimalist Program – the statute of which is not yet clear to me (and not even to the generativists themselves, so it seems to me).

²⁹ Translated from the Portuguese original by the author.

2.5. The Minimalist Program

I will start quoting an extract of the Presentation Eduardo Raposo wrote for the Portuguese translation of the book *The Minimalist Program* (Chomsky 1995):

*It is important to point out that the MP [Minimalist Program] is **not** a new theoretical framework of the transformational-generative grammar, in the sense the model P&P [Principle and Parameters], or the Extended Standard Theory, or the Standard Theory are. In this sense, the MP does not replace the P&P model. Quite the opposite, MP stands crucially on the P&P model, and starts from it to propose some new questions that could not, in fact, be conceived outside that model. In a way, MP is a set of 'guidelines' oriented by the intuitive idea of avoiding the postulation of theoretical entities that are not conceptually necessary inside the theory's logic. (Raposo 1999, pp. 15-16)³⁰*

In Raposo's words, MP should be considered not only part of the GG Program, like all other models we have seen so far, but also as part of the model that resulted from the last major elaboration of the program: the Principle and Parameters Theory.

Actually, MP would consist of just 'guidelines' of methodological nature to help the linguists to apply the 'Ockham's razor' to the Principles and Parameters Theory (P&P), eliminating what was unnecessary, basically for theoretical economy reasons³¹.

Being so, MP would not raise much interest in our investigation. We have been dealing with the major 'changes of course' of the Program (from LSLT to the Standard Theory and from the Standard Theory to P&P) as creative changes in the heuristics, resulting in a new theory. Considering that MP is not accepted as a creative change in the heuristics, resulting in a new theory, basically we have nothing to say about it.

However, the framework is not so clear and I believe we have to investigate more carefully the nature and function of the MP in the framework of the GG Program.

In order to understand the MP role, we must understand better the working hypothesis of human languages proposed by the GG program³².

According to Chomsky, languages are biological systems that men use to speak about the world (or about the mental representation they have about it), describe, refer to, ask, communicate with one another, articulate thoughts, talk to themselves, etc. Those 'things' we do with language constitute what Chomsky calls the *conceptual-intentional system*. On the other hand, as an 'expressive' medium, language must be associated to a production and reception system, of motor-sensorial nature, capable of allowing for the production and reception of sounds that constitute the linguistic expressions. Chomsky labels this second system *articulatory-perceptual system*.

Thus, the human language must be able to contact (be an interface of) not only the conceptual-intentional system (C-I) but also the articulatory-perceptual system (A-P)³³.

³⁰ Translated from the Portuguese original by the author.

³¹ There is a concern about showing that the human languages are themselves economical and perfect (I thank Evani Viotti for this observation).

³² In what follows, I make strong use of Raposo 1999.

³³ The question of interfaces is an old question. In the standard theory, it was up to the deep structure to make an interface with the C-I system, and to the surface structure to make the interface with the A-P system. In the government and binding theory the interface with the C-I system is done in Logical Form and it is up to the Phonological Form to make the interface with the A-P system. Certainly, although the idea is very old, the devices to actualize these interfaces have become more and more sophisticated as the time passes by.

According to Chomsky, the C-I and A-P systems have their own structure and are independent from the human language. It is possible to assume that they impose conditions over language. It is reasonable to think that human languages have the articulatory and hearing capacity of human beings as their limitations, for instance³⁴.

So, for us to use languages it is necessary that the linguistic expressions satisfy certain conditions imposed by these two outside systems. Thus, the MP fundamental question is the establishment of the ‘measure’ that allows the evaluation of the ‘optimality’ of the structures in satisfying the conditions imposed by the outside systems. In other words, it will be considered ‘good’, ‘grammatical’, ‘acceptable’ the structure that fully satisfies the interface conditions.

Now, though it can sound like big news, it seems that we are facing the same ‘movement’ that led GG to switch from rules to principles: that of obtaining the maximum generality with the least resources. In a way, following the strong assumption of ‘psychological reality’ of the computing systems – assumption that, in a higher or lower degree, has been guiding the generative analysis from their first formulations – what is being assumed is that the conditions over structures, realized in the auxiliary theories, are imposed by performance, by ‘pragmatics’, by the ‘use’ we make of them. And this ‘use’ involves not only the elements of comprehension of the world (the C-I system) but also sensory-motor elements (the A-P system). The bet made by the MP is that these general conditions, coming from performance, are capable of imposing proper conditions to the computing system in order that it works in an ‘optimal’ manner and be able to perform the task of the theories of the government and binding model with more economy, considering that it does not need to postulate anything but the interfaces: syntax is reduced to the minimum.

It is no longer necessary to postulate restrictions on structures. The ‘well-formed conditions’ of the structures, essential in the other stages of the program, disappear, and the guarantee that a structure is well formed (grammatical) will depend on the degree of satisfaction of the conditions imposed by the outside systems (C-I and A-P) that the structure presents. In other words, it will be considered more ‘adequate’, ‘acceptable’, ‘well-formed’, ‘grammatical’ that structure that best satisfies the phonetic production/reception and semantic conditions.

From this point of view, MP is nothing more that the radicalization of the movement that led the GG program to replace rules by principles.

3. Conclusion.

The history of GG shows three major ‘strategies’ in the delimitation of the language competence present in the speaker’s mind/brain. At time 1 (the SS Theory), grammar should generate the sentences of the language directly (into their surface structures). They dealt exclusively with syntax (maybe phono-morpho-syntax), and the notion of *generative grammar* was similar to the current notion in logics and mathematics.

At time 2 (the Standard Theory), grammar begins to generate abstract objects that are interpreted in the sentences of the language (in its phonetic form and in its meaning), that is, the set of abstract objects generated by the grammar is projected into the language, describing it as a set of possible signifiers linked to a set of possible meanings (pairs $\langle s, m \rangle$, where s is a signifier and m is a meaning). Here the notion of generative grammar undergoes a slight modification in relation to its previous meaning:

³⁴ The limit can also be *visual*, as long as we consider that sign languages of the deaf are also natural languages and actualize from the same biological matrix than the, say, ‘audio-oral’ languages.

it does not generate the sentences of the language directly. However, the commitment with the notion of *language* remains the same, since grammar will generate as many abstract objects as there are sentences in the language and no more. This makes it possible to go on considering grammar as a ‘*system of rules that generate the sentences - and only the sentences - of the language*’³⁵.

At time 3 (P&P), grammar generates abstract objects that explicit the properties that the speakers take into account when they issue evaluations on the grammaticality of *linguistic objects*. The sentences of any language constitute only a sub-set of this set of linguistic objects and so it is never - and under no criterion - possible to say that grammar generates the sentences of the language - at the most it is possible to say that grammar *allows (licensees)*, among other things, the sentences of a given language³⁶.

Using Lakatos’s terminology, we can say that such ‘strategies’ characterize different *heuristics* and that the GG program has experienced two major *creative shifts*: the first involving SS Theory and the Standard Theory, the second involving the Standard Theory and P&P. We can also consider that the best periodization of the development of the program establishes three periods: the SS period, which goes from Chomsky’s work (around 1954) to the publication of ‘Aspects’ (1965); the Standard Theory period, which starts with the publication of ‘Aspects’ and goes till ‘Conditions on Rules of Grammar’ (1976a)³⁷; and the P&P period, starting with Chomsky 1976a and continuing till today. In its first period, the theory is still very tied to the structuralist ways of doing linguistics and, consequently, there is a certain conflict between the requirements of the program and its theoretical availability: it is an unstable period in the theory. In the second period, it is the very program that faces difficulties: there is a proliferation of alternative heuristic proposals and there are dissenters. This second period is characterizes by disputes and by a great theoretical ebullience. The third period is a period marked by a great development in the expansion of the empirical content of the theory: a large number of languages are analyzed to a satisfactory degree and the principles are established in a very consistent form.

Finally, I believe that a few words about Noam Chomsky’s role in this story are in place.

Chomsky has always been the great leader of the Generative Community, imposing advances, rewriting the program, rejecting and/or supporting proposals. Chomsky acts - and he is seen like that by the community - as the ‘owner’ of the program, the person who utters the last words about the validity of the research lines proposed by his associates, the person who says what must and what must not be researched, the person who from time to time carries out a ‘balance’ of profits and losses (conquests and theoretical costs) of the theory and proposes the major syntheses

³⁵ The SS grammar had as output of the set of conversions the very sentences of the language; the standard theory had as its output of the set of conversions a set of structures (syntagmatic indicators) that were interpreted in the sentences of the language.

³⁶ In Chomsky’s words: ‘*We may perfectly well think of the grammar of, say, English, as assigning a structural description to every possible sound. Some will be characterized simply as noise, others as sounds of perhaps some language (but not mine), others as expressions of my language with some figurative interpretation, others as paired with strict “literal interpretations,” and so one.*’ (1981, p. 5)

³⁷ Usually we find references to the text ‘Conditions on Transformations’ (Chomsky 1973) as being the text that allows a ‘quality leap’ in the GG program. However, I believe that no matter how important this text might be in the process of change from a system of rules to a system of principles, in fact, it does not modify any of the fundamental concepts of the program. The text ‘Conditions on Rules of Grammar’ (Chomsky 1976a), on the other hand, because it releases the syntactic component, allowing for the ‘overgeneration’, forces Chomsky to recognize that the program will only have a chance if its notion of language is abandoned. In my opinion, it is this second text that opens the third period of the GG program. The fact the critics/commentators of Chomsky’s work do not recognize in ‘Conditions on Rules of Grammar’ the importance I do, make me feel uncomfortable, but I cannot be unfaithful to my convictions.

which will provide the new directions. Without much exaggeration, we could say that GG has always been an essentially Chomsky's creation. All those who did not agree with Chomsky, at some point in the history of the program, either surrendered to the power of the 'master', rejoining the 'good path', or became dissenters, remaining marginal to the program. No matter how interesting the proposals presented by Chomsky's associates were, they are only incorporated effectively to the program's theoretical arsenal after having been supported by Chomsky. Chomsky's centralizing power is so strong that it is possible to find a book of about 250 pages devoted exclusively to a survey of the destiny – invariably unhappy – of those who dared to defy him (Botha 1989).

Nevertheless, an interesting fact starts to unfold. The 'cold' reaction from the community to the Minimalist Program proposed by Chomsky seems to indicate that we are living a moment in which the program is reaching a certain maturity and that the community can already walk on their own feet, dispensing with Chomsky's custody. Any way, it is too early to make any reliable evaluation of the course the generative enterprise will take from here.

4. References:

- Anderson, S.; Kiparsky, P. (eds.) 1973. *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Winston.
- Bloomfield, L. 1926. A set of postulates for the science of language. *Language* 2, pp. 153-164.
- Borges Neto, J. 1991. *A Gramática Gerativa Transformacional: um ensaio de Filosofia da Lingüística*. Tese de doutorado, IEL-UNICAMP.
- Botha, R. 1989. *Challenging Chomsky: The Generative Garden Game*. New York: Basil Blackwell.
- Chomsky, N. 1951. *Morphophonemics of Modern Hebrew*. University of Pennsylvania (unpublished M.A. dissertation).
- Chomsky, N. 1953. Systems of syntactic analysis. *Journal of Symbolic Logic* 18 (3), pp. 242-256.
- Chomsky, N. 1955a. Logical syntax and semantics: their linguistic relevance. *Language* 31, pp. 36-45.
- Chomsky, N. 1955b. *Transformational Analysis*. University of Pennsylvania (unpublished Ph.D. thesis).
- Chomsky, N. 1955c [1975]. *The Logical Structure of Linguistic Theory*. New York: Plenum.
- Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.
- Chomsky, N. 1966. *Topics in the Theory of Generative Grammar*. The Hague: Mouton.
- Chomsky, N. 1967. Remarks on Nominalization. In Chomsky 1972 (Spanish translation in Sanchez de Zavala (comp.) 1974, pp. 133-187).
- Chomsky, N. 1968. Deep structure, surface structure and semantic interpretation. In Chomsky 1972 (Spanish translation in Sanchez de Zavala (comp.) 1974, pp. 276-334).
- Chomsky, N. 1971. Some empirical issues in the theory of transformational grammar. In Chomsky 1972 (Spanish translation in Sanchez de Zavala (comp.) 1974, pp. 444-524).
- Chomsky, N. 1972. *Studies on Semantics in Generative Grammar*. The Hague: Mouton.
- Chomsky, N. 1973. Conditions on Transformations. In Anderson & Kiparsky (eds.) 1973,

- pp. 232-286.
- Chomsky, N. 1976a. Conditions on rules of grammar. *Linguistic Analysis* 2, pp. 303-351.
- Chomsky, N. 1976b. On Wh-movement. In Culicover et. al. (eds.) 1977, pp. 71-132.
- Chomsky, N. 1977. *Dialogues avec Mitsou Ronat*. Paris: Flammarion (tradução para o português: *Diálogos com Mitsou Ronat*, São Paulo: Cultrix, s/d).
- Chomsky, N. 1981. On the representation of form and function. *The Linguistic Review* 1 (1), pp. 3-40.
- Chomsky, N. 1982a. *The Generative Enterprise*. Dordrecht: Foris.
- Chomsky, N. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press. (Portuguese translation: *O Programa Minimalista*, Lisboa: Caminho, 1999)
- Culicover, P.; Wasow, T.; Akmajian, A. (eds.) 1977. *Formal Syntax*. New York: Academic Press.
- Dascal, M. (org.) 1978-1982. *Fundamentos Metodológicos da Lingüística*, 4 volumes, São Paulo: Global (vol 1); Campinas: organizer's edition (vol. 2, 3 e 4).
- Dascal, M. 1978. As convulsões metodológicas da lingüística contemporânea. In Dascal (org.) 1978, vol. 1, pp. 15-41.
- Emonds, J. 1970. *Root and Structure Preserving Transformations*. Bloomington: Indiana University Linguistic Club (IULC).
- Feyerabend, P. 1974. "Consolations for the Specialist" in Lakatos & Musgrave (eds), p. 197-230.
- Galmiche, M. 1979. *Semântica Gerativa*. Lisboa: Presença.
- Harris, R.A. 1993. *The Linguistic Wars*. Oxford: Oxford University Press.
- Harris, Z.S. 1952. Discourse Analysis. *Language* 28, pp. 1-30 e 474-494.
- Harris, Z.S. 1957. Cooccurrence and transformation in linguistic structure. *Language* 33, pp. 293-340.
- Harris, Z.S. 1965. Transformational Theory. *Language* 41, pp. 363-401.
- Harris, Z.S. 1968. *Mathematical Structure of Language*. New York: Intersciences.
- Hockett, C.F. 1955. A manual of phonology. In *Antropology and Linguistics: Memoire II*. Baltimore: Indiana University Publications.
- Kato, M. 1974. *A Semântica Gerativa e o Artigo Definido*. São Paulo: Ática.
- Lakatos, I. 1970 Falsification and the methodology of scientific research programmes. In Lakatos & Musgrave (eds) 1974, p. 91-196 (republished in Lakatos 1978 p. 8-101).
- Lakatos, I. 1971. History of Science and its Rational Reconstructions. In Lakatos 1978, pp. 102-138.
- Lakatos, I. 1978. *The Methodology of Scientific Research Programmes* (Philosophical Papers vol 1). Edited by John Worrall & Gregory Currie. Cambridge: Cambridge University Press.
- Lakatos, I. & Musgrave, A. (eds) 1974. *Criticism and the Growth of Knowledge*. (3rd impression) London: Cambridge University Press.
- Lakoff, G. 1971. On generative semantics. In Steinberg & Jakobovits (eds.) 1971, pp. 232-296.
- Lees, R. 1957. Review of Noam Chomsky, *Syntactic Structures*. *Language* 33, pp. 375-408 (reprinted in Otero, C. (ed) 1994. *Noam Chomsky: critical assesment*. Vol 1 London; Routledge).
- Lobato, L.M.P. 1986. *Sintaxe Gerativa do Português: da teoria-padrão à teoria de regência e ligação*. Belo Horizonte: Vigília.
- Milner, J-C. 1973. Écoles de Cambridge et de Pennsylvanie: deux théories de la transformation. *Langages* 29, pp. 98-117.

- Newmeyer, F.J. 1980. *Linguistic Theory in America: the first quarter century of transformational generative grammar*. New York: Academic Press
- Peters, S.; Ritchie, R. 1969. A note on the universal base hypothesis. *Journal of linguistics* 5, pp. 150-152.
- Peters, S.; Ritchie, R. 1971. On restricting the base component of transformational grammars. *Information and Control* 18, pp. 483-501.
- Peters, S.; Ritchie, R. 1973. On the generative power of transformational grammars. *Information Sciences* 6, pp. 49-83.
- Quicoli, A.C. 1976. Conditions on clitic movement in Portuguese. *Linguistic Analysis* 2 (3), pp. 199-223.
- Raposo, E.P. 1999. Da teoria de Princípios e Parâmetros ao Programa Minimalista: algumas idéias-chave. Apresentação da tradução portuguesa de Chomsky 1995, pp. 15-37.
- Ross, J.R. 1967. *Constraints on variables in syntax*. Bloomington: Indiana University Linguistic Club (IULC).
- Sanchez de Zavala, V. (comp.) 1974. *Semántica y sintaxis en la lingüística transformatoria I*. Madri: Alianza.
- Wells, R. 1947. Immediate Constituents. *Language* 36, pp. 360-367.