Theories of language

Jan Koster

1. Introduction

Since Antiquity, a central concern of theories of language has been the question whether language is predominantly a matter of "nature" or of "nurture." One version of this dilemma is whether language is primarily a socio-cultural reality or a biological phenomenon. British Empiricism and German Romantic ideas, interacting in complicated ways, set the stage for much of 19th century linguistic thinking, which culminated in the various structuralist schools of the first half of the 20th century. Often, this tradition emphasized culture, nurture and diversity. In the second half of the 20th century, influenced by Chomsky, nativism and the idea of Universal Grammar made a powerful come-back. This culminated in the "biolinguistic" idea that language has a core that can be compared to an organ, or rather, to a computational system of the kind found in mammalian vision. Instead of embarking upon the impossible task of giving an overview of all current theories of language, I will give a historical sketch of how the Chomskyan-style linguistics fared with respect to the perennial tension between culture and biology in the study of language and how this tension can be resolved with current neurobiology. It is my hope that this story suggests some morals about other theories of language as well.

2. The revolution

During the second half of the 20th century, then, linguistic theorizing was dominated by the so-called Chomskyan revolution. This type of linguistics rose to ascendancy with Chomsky (1957), had its greatest popularity during the 1960s, culminated in the 1970s, but steadily lost ground after the 1980s. At the time, said development was generally seen as revolutionary, although there were always critics and skeptics. It is questionable whether the theories in question were as revolutionary as believed at first. In retrospect, a good case can be made that early generative grammar, rather than superseding older theories, was largely ignorant of them. Let me explain.

Pre-Chomskyan linguistics, at least in Europe, was dominated by the Saussurian idea that language is primarily a system of signs, of which words (morphemes) are the most important (Saussure 1916). Signs were thought to have a public face (*signifiant*) and a conceptual side (*signifié*). In the most common case (but not necessarily so) the public face of a linguistic sign is formed from the sounds of speech. But also the conceptual side of language was believed to have a public aspect, as it was assumed that the concepts of a language represented particular choices from a perhaps universal but more or less amorphic conceptual "space." Conceived this way, language was seen first and foremost as a social-cultural reality. Of course, this did not exclude that such social-cultural realities were possible only on the basis of psychological or biological capacities.

Next, signs (words) were believed to enter into paradigmatic and syntagmatic relations. If we limit ourselves to syntax, we can say that paradigmatic relations define a class of elements that can take the same position in a sentence, such as *John* and *The father of John* in (1):

(1) a. *John* left
b. *The father of John* left

Such a paradigm is also called a substitution class or a distribution class (in American structuralism).

Syntagmatic relations are the horizontal relations in a phrase or clause, like the relation between *John* or *The father of John* with the following verb *left* in (1). It was generally recognized that the syntagmatic relations of language are not between single words but between *groups of words* (also known as phrases or constituents). As a result, sentences were analyzed as having a hierarchical structure. It was also recognized that parts of phrases could be self-similar, meaning that noun phrases could contain noun phrases or clauses a clause. In the later Chomskyan tradition, under the influence of a pre-occupation with formal languages and recursive-function theory, this self-similarity was generally referred to as "recursion." Both phrase structure and recursion, then, were, terminology aside, within the scope of pre-Chomskyan structuralism.

It should be noted, in the light of later claims to the contrary, that pre-Chomskyan theories of phrase structure were not necessarily construction-bound: the notion of a word group or phrase is more abstract than the notion of a construction (active, passive, question, etc.). Both active and passive constructions, for instance, were analyzed in terms of word groups built up from the same ingredients (such as heads, complements and adjuncts). More generally, word groups (phrases, constituents) were seen to express the combinatorial potential of a word (the head or the core of the group). This potential was often referred to as the "valency" of the word (Tesnière 1959), which was ultimately believed to be a reflection of the semantic properties of the word. Using somewhat anachronistic terminology, it could be said that sentence structure was seen as hierarchical and recursive, consisting of word groups projected from the lexicon. The items of the lexicon (Saussurian "signs") were held to be socio-cultural objects and part of the inventory referred to as "langue." As syntactic structures were conceived of as the realizations of the valency of words, syntax was socio-cultural in orientation as well. Matters of individual psychology ("parole") were only secondary and reduction to biology was practically unheard of.

Phrase structure theory was only one element of structural linguistics. Another aspect was analysis in terms of information structure as developed by the Prague School in the 1920s and 1930s. According to that type of theory, the content of a sentence can be divided in old and new information (theme-rheme, topic-comment, focus-presupposition). Analysis in terms of information structure is also implicit in the old "Satzklammer" theory, which divides the sentence in a "middle field" and a "left" and "right periphery" (see Kathol 2000). Both word order variation on the middle field and specific use of the peripheries was seen in terms of what we now would call information structure. Consider, for instance, the fact that a sentence consists of a DP and VP. In most languages of the world, the DP precedes the VP, reflecting the fact that it is a natural organization of information that the topic of a sentence (the DP) is mentioned before something is said about it (in the VP). Especially with respect to linear order, information structure was seen as a leading principle.

What is known as the Chomskyan revolution introduced some technical innovations and sought to give the field a new metatheoretical orientation. In *Syntactic Structures* (Chomsky 1957) a language L is a set of sentences generated by the grammar. The technical notion "grammar" was algorithmic and inspired by the study of formal languages and recursive function theory. More precisely, the sentences of L were generated by context-free phrase structure rules followed by several modifications generated by transformational rules. Central in this endeavor was the notion of a "kernel sentence," based on the intuition that simple, declarative sentences are somehow basic. The strings underlying kernel sentences (called "deep structure" in Chomsky 1965) were generated by the phrase structure rules, whereas the kernel sentences themselves were generated by obligatory transformations (like

the rules accounting for subject-verb agreement). Other sentence types, like passive sentences and questions, were derived from (the strings underlying) kernel sentences by optional transformations. Sentence recursion was accounted for by so-called "generalized transformations," which combined phrase markers (the structures resulting from phrase structure rules and "singular" transformations) into complex structures.

According to the new metatheoretical orientation, what was formalized this way was not a socio-cultural reality but a matter of individual psychology (ultimately biology), referred to as "competence." All of this had great appeal at the time, as it suggested explicitness and mathematical sophistication the field had not seen before.

Unfortunately, it all rested on shaky grounds, as became clear in the next 15 years. The failure was not recognized as such but perceived by many generative grammarians as continuous innovation. Although transformational grammar died in the 1970s, the idea that there was a major revolution in linguistics between 1955 and 1970 is kept alive until the present day. In fact, what came after 1980 is even believed by some to be a second revolution, while it was, in spite of some minimal innovations, in reality the beginning of a period of theoretical stagnation that has been with us ever since.

What was problematic about transformational-generative grammar from the beginning was insufficient understanding of the fact that the formal languages used were an extremely poor tool to model natural language. First of all, to the extent that formal languages define a set of well-formed formulas, there is no close natural language (NL) equivalent. Strings of NL words are more or less interpretable, no matter their grammatical status. Well-formedness in NL is a relative matter and not something absolute. Compare the following two strings:

(2) a. John read the book
b. John book read

For a native speaker of English, (2a) is fully well-formed, while (2b) is not. Nevertheless, (2b) is clearly interpretable and might even receive the same interpretation as (2a). When English is defined as a set of sentences L, there is no non-arbitrary reason to include (2a) in that set and to exclude (2b). Of course, (2a) is "optimal" in a sense that (2b) is not, but the point is that there is no similar notion of "optimality" with respect to artificial formal languages. In the latter, a string is well-formed or not. To the extent that artificial languages are interpreted by some model-theoretic semantics, this semantic interpretation only applies to the well-formed formulas. The set of semantically interpretable NL strings is far too complex to be defined by any sensible algorithm.

Although this point was appreciated early in the history of generative grammar, the obvious conclusion was hardly drawn: a NL cannot be satisfactorily characterized by an algorithm. In current minimalist theories, semantic interpretation –blocked by "uninterpretable" features at one of the grammar's interfaces– even has the magic power to let suboptimal derivations "crash." However, as (2b) and numerous other examples may show, semantic interpretation does not have such powers but tries to make the best of any string of words, no matter its degree of optimality from a syntactic point of view.

There is another difference between formal languages and NL that played a much bigger role in the subsequent development of generative grammar. Unlike the symbols manipulated by the rules of formal languages, NL words have rich inherent properties independent of the syntactic rules in which they appear. First of all, the words of NL have a public side ("signifiant") and a conceptual side ("signifié") and are, in that sense, complete interface elements, even *before* entering into derivations that will be sent at the end of some "phase" to the interfaces involving sound and meaning. Furthermore, all words have a valency, predicting in which syntactic environments they may occur. Thus, the English word

book may be preceded by an indefinite article ("a book") rather than be followed by it (*"book a"). On its right side, the noun may be followed by a PP ("a book about linguistics"), etc. So, each word comes with a set of potential environments that can be seen as a property of that word. The syntactic environments are partially predictable on the basis of the conceptual structure associated with a word. No such properties are associated with the symbols of formal languages, which have no properties beyond the fact that they can enter into certain rules.

Not only the valency of words was practically ignored in Syntactic Structures, so was the information structure discussed by the Prague School. Thus, a rule like S → NP VP is arbitrary in that it could just as well be S → VP NP, ignoring the fact that in the majority of natural languages, the subject precedes the predicate.

The new metatheory was developed partially in response to the behaviorism of Skinner and Quine (Chomsky 1959). It reinterpreted the socio-cultural "langue" as "competence," a matter of individual psychology, to be ultimately explained in terms of biology. This move had a minimum of plausibility thanks to the shift from word-oriented linguistics to syntax-oriented linguistics. Starting from signs (morphemes, words), claimed to be the right approach in this article, it is immediately clear that the reconstruction of syntax solely in terms of individual psychology is incoherent. Words, obviously, are collective property, not belonging to an individual but to a speech community.

3. Rediscovering the lexicon

Transformational-generative grammar received its classical form in *Aspects of the Theory of Syntax* (Chomsky 1965). According to the Aspects model (also known at the time as the "standard theory" or ST), a grammar consists of a context-free phrase structure grammar and a lexicon (the "base component") responsible for the "deep structure," which is the input for semantic interpretation. Deep structures could be modified by a series of cyclically organized transformations leading to "surface structures." The generalized transformations disappeared in this model, with recursion now being a property of the base component. This development followed the lead of Katz and Postal (1964), who had hypothesized that all semantic interpretation depends on deep structure. This trend was continued by what became known as Generative Semantics, which held that 1) syntax is partially pre-lexical, manipulating concepts rather than morphemes or words, and 2) that the inventory of categories at the deepest level could be limited to S (proposition), V (predicate) and NP (argument). For many this suggested a simplified universal syntax inspired by formal languages once more, in this case the very poor syntax (compared to NL syntax) of predicate logic. Many linguists did not find this empirically plausible in the long run and it is fair to say that Generative Semantics died out in the early 1970s.

Chomsky himself gave up the idea that semantic interpretation solely depends on deep structure and demonstrated that transformations, particularly those modifying linear order, affected semantic interpretation (Chomsky 1971). This led to the Extended Standard Theory (EST), which dominated the field until the end of the 1970s, when the emphasis shifted to "principles and parameters" (Chomsky 1981, also referred to as the Government-Binding theory or GB).

The effects of linear order on semantic interpretation reminded of the Prague-style concern about information structure and it was the first time transformational-generative grammar "interfaced" with this important structuralist tradition.

Much more important is the obvious but underexposed fact that transformational-generative grammar actually did not survive the introduction of a lexicon in *Aspects*. Within

five years, it became (or should have become) clear that the ensuing lexicalism was not a continuation of a revolution but a return to the basic tenets of pre-Chomskyan structuralism (particularly in some of its European forms). First of all, it became clear that context-free phrase structure rules become superfluous as soon as a lexicon with subcategorization frames is introduced. In Aspects, syntactic base structures are generated largely independently of the lexicon, whereas lexical items are inserted later in the derivation on the basis, among other things, of a match between the subcategorization frame and a structure generated by the phrase structure rules. In the following example (simplified but in the spirit of *Aspects*), a VP is expanded as in (3a) while the verb *see* can be inserted on the basis of this verb's subcategorization frame given by the lexicon (3b):

(3) a. VP → V NP
b. *see*: [+V, --NP]

If things are set up this way, the same information is given twice: (3a) gives the structure of a common VP-type and this structure is stated once more in the subcategorization frame (3b). Up until the present day, there is only one plausible way to get rid of this redundancy, namely by giving up (3a) altogether and by "projecting" the structure from the subcategorization frame (3b) instead (but see Maling 1983). Another way to put this is to say that in order to generate a VP for a verb like *see*, it suffices to spell out the valency of *see*, as specified in (3b). The redundancy was in principle eliminated by the rediscovery of X-bar theory in Chomsky (1970), which was also motivated by the need to account for endocentricity. The value of X-bar theory in solving the redundancy problem was confirmed in Chomsky (1981: 31).

4. Not quite a revolution

Unfortunately, the implications of giving up context-free phrase structure rules (like 3a) in favor of projection from lexical frames (like 3b) were underappreciated at the time. What was new about generative grammar in the 1950s was not something like (3b) but description in terms of rules like (3a), which were adapted from the study of formal languages and mathematical linguistics. Very few pre-Chomskyan structuralists would have objected to the idea that syntax spells out the valency of lexical items. It is therefore reasonable to say that giving up (3a) in favor of projection from the lexicon was not the next step in an ongoing revolution, but reinventing the wheel of structuralist grammar. Early generative grammar was not a revolution but a development based on the serious error that NL syntax can be largely developed independent of the lexicon.

What has been appreciated even less, until the present day, was that the shift from a socio-cultural view of language to individual psychology was based on the same error. Obviously, lexical items are not a matter of individual psychology but artifacts belonging to a shared culture. So, also in this respect, the adoption of lexicalism was an implicit return to a major tenet of Saussurian structuralism.

What about the rest of the revolution? Phrase structure grammar, although replaced again by pre-revolutionary projection of lexical valency, was never the most controversial part of transformational-generative grammar. More controversial was the introduction of transformations, together with the claim of their "psychological reality." The same can be said about the closely related distinction between deep structure and surface structure. The meta-theory not only made claims about the individual-psychological nature of grammars, it was also hypothesized that in language acquisition, grammars are selected from a small class of

feasible (or possible) grammars. This innate hypothesis space was referred to as LAD (the Language Acquisition Device) or Universal Grammar. In its most extreme form, UG was seen as a single grammar with parameters, to be set on the basis of the data the child was exposed to during the acquisition of his or her native grammar.

I will come back to how the lexicalism of around 1970 should have had affected said theory of language acquisition. At this point, I will limit myself to the question how the other two "revolutionary" tenets fared in the 1970s: the idea of transformational analysis and the idea of grammar with multiple levels (like deep and surface structure). To make a long story short, both ideas appeared to be mistaken. One transformation after another disappeared, as in many cases a classical transformational analysis was shown to be impossible (see for instance Jackendoff 1969, Wasow 1972 and Higgins 1973). In most cases, transformations were replaced by construal rules, which can be seen as rules of completion based on feature sharing (see Koster 1987: 8). Consider a typical reflexive, like *himself* in (4):

(4) *John* saw *himself*

In earlier transformational analyses, a transformation derived (4) from *John saw John*, replacing the second *John* by *himself*. *John saw John* was supposed to be closer to the deep structure and seen as a *full reconstruction* of the input level for semantic interpretation. Full reconstruction did not work, for instance for sentences with quantifiers: *everyone saw everyone* does not mean the same as *everyone saw himself*. So, it was concluded that a sentence like (4) only involved *partial reconstruction*: the anaphor *himself* was directly introduced at the deepest level and the only completion provided by further rules was the referential identity shared with *John* (commonly expressed by a referential index). Full reconstruction was the core idea of transformational analysis, going back to Zellig Harris' notion of normalization.

In most cases, partial reconstruction (by "construal") was accepted, but for reasons that have remained obscure until the present day, an exception was made for displacements, i.e., the structures derived by movement transformations. The most typical of those are NP-movement (as in 5a) and Wh-movement (as in 5b):

(5) a. *John* was seen [$_{NP}$ --]
b. *What* did John see [$_{NP}$ --]

The empty positions [$_{NP}$ --] were referred to as "traces" at the time, but that was a notion that already presupposes displacement by movement rules. The real point is that elements like [$_{NP}$ --] express the fact that the valency of a verb like *see* remains intact under displacement. This opened the door for non-transformational analysis of displacement, i.e. in terms of partial reconstruction instead of full reconstruction. In structures like (5), the reconstruction is partial because, by property sharing between antecedent and dependent element (as in 4), the object positions only derive their lexical content from the antecedent. Their categorical status is given by the subcategorization frame of *see* in (5) (also involving theta-roles and Case), hence making the reconstruction partial instead of total.

The partial-reconstruction view of displacement makes it more similar to other forms of construal, like those for anaphoric relations and agreement phenomena. It made it possible to get rid of transformations altogether, as "movement" was the last stronghold of transformational analysis. Without transformations, X-bar theory sufficed for purposes of structure building, with uniform construal as a form of secondary computation. Important empirical confirmation of this view was provided by research in the wake of Emonds' (1970) structure-preserving hypothesis, which tended to show that the major movement

transformations did not create any structure beyond what was given by the base structures (of X-bar theory). The implications were dramatic but often ignored at the time. With X-bar theory as the main device of sentence generation and transformational analysis gone, the characteristic multiple-level approach (with deep and surface structure) was gone as well, leaving generative grammar as a non-revolutionary elaboration of the more traditional ideas of structuralist syntax. The conclusion that grammar is mono-stratal (i.e., with one level of representation for phrase structure) was drawn by some (see Brame 1978, Koster 1978, 1987 and Pollard and Sag 1994), but rejected by the mainstream. The refusal to draw this conclusion in the 1970s was partially responsible for the decline of Chomskyan generative grammar as a *theoretical* enterprise.

Movement lived on as "move alpha" and continued to be seen as the basis of a multiple-level theory in Chomsky (1981). In fact, things became even worse than around 1970 because an extra level of representation was introduced known as Logical Form (LF). LF was supposed to be derived by "invisible" movement (QR in May 1977, 1985). LF in this sense has always remained an *ad hoc* addition, as it fails to pass the only empirical test that could have made it interesting: conforming to the same locality constraints characterizing "overt" movement rules (seen as partial construal in the sense discussed above).[1] Overt displacement has a clear function, namely highlighting certain categories. Highlighting, often in the spirit of Prague-style information structure, crucially depends on the *visible* effects of displacements. Invisible movement does not make sense from this point of view. In early Minimalism, this fact was masked by giving "movement" an entirely new rationale, namely in terms of "feature checking." Feature checking was further divided between the checking of "strong" features for overt movement and "weak" features for covert (LF-)movement. All of this was arbitrary and *ad hoc*, unlike the earlier rationale for displacement in terms of information structure. Altogether, this shift to arbitrary feature checking was another symptom of the decline of mainstream generative grammar.

5. A partial consensus

In current Minimalism, the idea of multiple levels of representation (apart from the interfaces) is mostly given up, making mono-stratal syntax the consensus view now, more than 30 years after the demise of transformational analysis. Unfortunately, the full-reconstruction view of displacement is still maintained in current Minimalism in the form of "internal merge" and the so-called copying theory of movement. The plausibility of this bizarre theory is not helped much by the fact that it is proclaimed to be "perfect" (rejected in Koster 2009b). I will discuss further problems of Minimalism in a moment.

Although we have latter day Minimalism and anti-lexicalist frameworks, like the theories based on distributed morphology, many linguists subscribe implicitly or explicitly to the lexicalist frameworks of the 1970s. The idea that syntactic structures, with minor additions, spell out the valency of lexical items completely undermined the claim that linguistics underwent a major revolution in the second half of the 20th century, as X-bar theory is conceptually not more than a variant of the phrase structure theories existing before Chomsky. This is not to deny that linguistics saw a spectacular growth in the last several decades, both in depth of analysis and empirical scope. There also is much more (and global) uniformity in terminology and means of representation than what was usual in the very disconnected forms of structural linguistics seen before the 1950s. Another change of lasting

[1] The island behavior of Wh-elements *in situ* in languages like Chinese or Japanese is not quite the same as what we see in the overt displacement relation and can be easily accounted for by the association of the Wh-phrase (*in situ*) with a scope marker.

importance was a methodological lesson learned from the natural sciences: *apparent* diversity of language data does not immediately falsify universal claims. What is considered "the facts" involves theoretical interpretation that can be just as wrong as the theory explaining the facts. This lesson seems to be lost on a growing number of linguists who think that theories of Universal Grammar can be falsified by just listing problematical data from a variety of languages (see, for instance, Evans and Levinson 2009 and the anti-universalist rhetoric of Tomasello 2008).

As a theoretical paradigm shift, however, early transformational grammar was much overhyped and practically everything that was radically new about it in the 1950s and 1960s appeared to be mistaken and was gradually given up after 1970. Within the (in retrospect) rather traditional lexicalist framework of X-bar theory, some earlier ideas were successfully further developed, like the greater use of empty categories (as $[_{NP}$ --] in (5) and the "silent" categories of Kayne 2005) and the addition of functional elements (like C in Syntactic Structures and INFL in later theories). This led to substantially "deeper" analyses than was common before, especially since the 1980s, when functional categories were hypothesized to project like lexical categories (DP, CP, IP, TP, AgrP, vP, etc.). Furthermore, the idea of "binary branching" was more commonly adopted than in earlier forms of structuralist syntax (Kayne 1984). It is important to emphasize that all these extensions can be adhered to without assuming a conceptual break or paradigm shift in the 1950s and 1960s.

The same is true for the one area in which generative grammar really shone. What I have in mind are the constraints on what was called "secondary computation" above. The most important of these are known as "locality principles." The three main types of locality principles are: 1) minimal distance principles, 2) minimal domain principles, and 3) command principles (Klima 1964, Reinhart 1976). Minimal distance principles go back to Rosenbaum (1967) and were further developed by Culicover and Wilkins (1984), Koster (1978) and Rizzi (1990). Minimal domain principles go back tot the A-over-A Principle (Chomsky 1964), which was famously criticized and developed in the form of island conditions (Ross 1967). Proposals to formulate simpler and more unifying versions of these conditions were Subjacency (Chomsky 1973), the Head Condition (Fiengo 1974, Van Riemsdijk 1978), the Bounding Condition (Koster 1978, 1987) and several others.[2] Island conditions (locality principles) are the core achievement of generative grammar, the main principles in what became known as the "principles and parameters" framework. In their standard formulation, these principles are not entirely construction-independent, as Subjacency (for movement constructions) has a form rather different from Principle A of the binding theory (of Chomsky 1981), but, at least in principle, a construction-independent account seems possible (cf. Koster 1987). Since X-bar theory is construction-independent as well, there clearly is justification for a construction-independent perspective next to approaches that emphasize the "Gestalt" of constructions (Goldberg 1995, Croft 2001).

All in all, then, before Minimalism, there was a growing generative consensus that (hierarchical and recursive) syntactic structures were projected from the lexicon to avoid the redundancy problem of the misguided forms of generative grammar of before 1970. Lexicon-driven sentence generation, we concluded, is not a revolutionary innovation of mid-20th century generative linguistics but an elaboration of the structuralist idea that syntax reflects the valency of lexical items. Many variants of generative linguistics (LFG, HPSG, etc.) interpreted this kind of grammar as mono-stratal, i.e., without multiple levels of representation like deep structure and surface structure. Since Minimalism and Construction Grammar adopted the idea of mono-stratal representation as well, there is nowadays near-consensus on this point (see also Culicover and Jackendoff 2005 for many similar

[2] See Koster (2007) for the elimination of variables from locality principles.

obervations). Minimalism, in practice, was a partial return to the idea of lexicon-independent sentence generation. In the next section, we will see if this move was justified.

6. Minimalism and the biolinguistic program

Since Chomsky (1995), Minimalism is one of the mainstream frameworks in current generative grammar. So far, it has led to much empirical research but to little theoretical progress. My concerns are not about Minimalism as a program. On the contrary, I subscribe to the overall goal to construct a theory that makes grammar look as perfect as possible and that relegates as much as it can to "third factor" principles. My dissatisfaction is about how this program is carried out in practice. It seems to me that little *theoretical* progress has been made since the 1980s. Part of the theoretical stagnation is due to the fact that some key problems of earlier versions of generative grammar are either unresolved or ignored. But there are deeper problems, it seems, that involve the very foundations of the field.

It is impossible to completely review minimalist practice here, so, I will give just a few examples. I will briefly criticize the use of Merge, a core notion of the theory, in a minute. Internal Merge is the successor of "move alpha" in earlier theories and just as problematic as "movement." Not only is the long tradition of criticism of such notions largely ignored, Internal Merge is even declared to be "perfect" these days, on very shaky grounds (see Koster 2009b for discussion). Similar things can be said about the core principles of the theory, the locality principles. It is occasionally observed that the Chomsky's current Phase Impenetrability Condition is very much like the Head Condition in Van Riemsdijk (1978) (see Boeckx and Grohmann 2004). What remains unsaid, however, is that the earlier condition was claimed to be better in the lively debate of 30 years ago because it was also involving the PP. If such issues are not forgotten, they are hardly mentioned anymore. It is difficult to see any progress here and the conclusion seems justified that no substantial progress has been made as to the nature of locality principles since the 1980s.

This is not a small point because, as we mentioned before, locality principles are by far the most important constraints of the theory that likes to call itself a "principles and parameters" theory. As for parameters, practice has perhaps even been less fruitful. The notion "parameter" has hardly been developed beyond the traditional observation that there are "differences" among languages, for example, as regards pro-drop or the order of head and complement. In short, the interesting principles were mostly discovered before Minimalism and the notion "parameter" has always remained underdeveloped from a theoretical point of view.

More generally, current generative grammar is often referred to as "a computational theory," but the current style of the field is so informal that it is practically impossible to find explicit accounts of what exactly is computed and how. With the low level of explicitness considered normal these days, references to "recursive Merge" say little beyond the traditional wisdom that syntax involves hierarchical structures with phrases within phrases of the same kind. There is nothing wrong with that insight but it would be somewhat exaggerated to say that it is revolutionary.

In order to see what is problematic about standard generative grammar, we must have a closer look at its foundations. During most of the life cycle of standard generative grammar, it was realistically interpreted in terms of "individual psychology," although it is true that from the beginning, a connection was made with biology, as in Lenneberg (1967). However, it is only recently that biological terminology became more dominant than references to "psychological reality," particularly since Jenkins (2000) and Hauser, Chomsky and Fitch (2002).

Note that neither rationalism nor the biological foundations of language in general are at issue here. I assume that all learning is constrained by biological principles, some of them very specific. The days of behaviorism are far behind us indeed. I consider the view of biologically constrained learning trivially true and from this general perspective language is as biologically based as our ability to play chess or to drive a car. None of that can be done by other organisms after all and at best there is a question as to how specialized or inflexible parts of our brain are with respect to certain tasks. So, the claim that language makes use of biological, innate components is obviously true and therefore trivial as far as I am concerned. This, of course, does not mean that it is a trivial task to find out what is the exact nature of the biological components involved.

The non-trivial claim of current biolinguistics is more specific, namely that grammar is like a specialized organ, or rather like internal computational systems such as the ones found in the mammalian visual system.

The trivial claim is true on general grounds and the non-trivial claim is false, or so, at least, I will argue. This has to do with the meaningful distinction that can be made between biologically-based functionality and culturally-based functionality. Consider, as an example, the role of the lungs in respiration and the role of the lungs in the playing of a wind instrument like the trumpet. The former is based on our genetic program and grows automatically, without human interference. Playing a trumpet, however, is based on the same innate structures –the lungs– but this time the function of the lungs is not formed by our genetic program but by coupling them with an instrument invented by humans in order to achieve human goals. The study of respiration in mammals is generally seen as biology, while playing the trumpet is generally seen as culture. This example illustrates once more that innateness is not the issue. The difference is based on the question whether the use of the same innate structures is mediated by human agency and artifacts (trumpet) or not (respiration).

So, the criterion distinguishing the trivial claim from the non-trivial claim is the involvement of artifacts. Language is only possible thanks to artifacts, namely our invented words. This simple observation suffices to refute the non-trivial claim. Whoever rejects this conclusion can choose between two ways out: either it must be shown that organs or internal computational systems (like in mammalian vision) also involve artifacts or it must be shown that the words of our language are not really artifacts. Clearly, the first option is absurd: neither organs like the heart or the kidneys nor mammalian visual computation is based on human inventions. Personally, I think the second option is untenable, too, but in practice it comes closest to a thesis defended in a certain variety of "biolinguistics." But let us first see how the internalist perspective was developed in generative grammar.

Although the internalist perspective was part of generative grammar since the 1950s, it has been characterized with the help of its current terminology since the early 1980s. Particularly in Chomsky (1986), the notions E-language and I-language were discussed, with the further argument that the proper subject matter of linguistics theory (in some reasonably narrow sense) is I-language. The distinction arose in opposition to Quine, who had claimed that languages are characterized by sets of well-formed formulas and that there is no issue as to what is the correct grammar as long as they generate the same set of sentences. Chomsky rightly rejected this view, characterizing enumerated sets of sentences as E-languages, while the actual mechanisms where characterized as I-languages, objects ("grammars") selected from a narrow range by children learning their language. The "I" in I-language stands for "internal" (to the human mind) and "individual" (also for "intensional," a notion to be ignored in this article).

The mistaken preoccupation with E-languages was attributed to the bias created by formal languages, as briefly discussed above. We also mentioned that the shift to I-language

led to incoherence, as there is no language without "public" words, which can, by definition, not be reduced to something strictly internal and individual. We also discussed the fact that the necessary introduction of a lexicon in *Aspects* (Chomsky 1965) led to a serious redundancy problem, that was only resolved by giving up lexicon-independent sentence generation and reinterpreting sentence structure as the spelling out the properties of words or morphemes (X-bar theory).

Up until the present day, it seems to me, this is the right way to do things. This does not mean, of course, that the specific form of the X-bar theory in Chomsky (1970, 1981) was correct. Like with any other empirical theory, one might hope for steady improvements over the years. In the case of X-bar theory, several modifications were proposed, for instance concerning the superfluousness of intermediate bar levels (Muysken 1982), the number of Specs or the nature of functional projections. I am not committed to any particular form of X-bar theory, only to the principle that redundancy can only be avoided by projecting syntactic structure straight from the lexicon. I see that insight as the firm conclusion of the first 25 years of generative grammar, even if it means a substantial return to the pre-generative tradition, when syntax was still word-oriented and not seen through the lens of formal languages with their lexicon-independent syntax.

Unfortunately, the introduction of bare phrase structure and Merge in Minimalism has compromised these insights. Bare phrase structure, as I see it, is a matter of notation. It is not substantially different from the versions of X-bar theory that gave up intermediate bar levels or the single-Spec restriction. Unfortunately, the adoption of Merge *as a sentence-generating rule* reintroduced the redundancy problem. What it comes down to is that Merge largely mimics what is already entailed by the subcategorization properties (the "valency") of lexical items. Both introduce the hierarchical, recursive and binary nature of syntactic structure, showing that a generalization is missed. All executions of Merge that I have seen assume the existence of subcategorization features. Chomsky (2000), for instance, claims that if two elements are merged, a feature F of one of the two elements must be satisfied. This element is called "the selector." Similarly, Hornstein, Nunes and Grohmann (2005) use subcategorization information for labeling purposes, where labeling mimics the information (head-complement structure) that was used by X-bar theory to account for endocentricity. But even label-free theories, like the one of Collins (2002), crucially depend on subcategorization. All of this reveals the obvious: even at the most elementary syntactic level (preceding the conceptual-intentional level according to Minimalism), Merge is not sufficient but can only account for endocentricity in combination with lexical information (subcategorization, valency).

In spite of the fact that in practice Merge only exists in tandem with said lexical information, the "recursive Merge only" thesis of Hauser, Chomsky and Fitch (2002), has the flavor of a theory that sees Merge as a *linguistic* computational mechanism completely independent of the properties of our invented words. This is perhaps even necessary for the non-trivial biolinguistics program to make sense. Lexicon-independent computation of this kind evokes reminiscences of the times (pre-1970) when natural language was still approached by many through the lens of formal languages. Lexicon-independent syntax has been refuted since the 1970s, largely on the basis of the redundancy problem. It is therefore hard to understand why it makes a partial comeback in minimalist theory and its biolinguistic interpretation.

The confusion can be taken away by seeing Merge not as part of a sentence-generating mechanism but as a meta-theoretical characterization of the properties of subcategorization frames: all involve binary hierarchical structure with the possibility of recursion. Sentence generation can remain to be seen as the spelling out of subcategorization frames, as in X-bar theory. In a word-oriented syntax of this kind, whatever Merge stands for is *applied* in the

creation of lexical items with complex combinatorial properties. In other words, there is a crucial and theoretically important difference between sentence generation *by* Merge and sentence generation as the spelling out of lexical properties *in accordance with* Merge. The former reintroduces the redundancy problem, the latter solves it.

In its application to lexical structures, then, Merge is not something biological but applied biology at best. Of course it is possible to consider Merge in abstraction of its lexical application and to reflect on its biological sources, but at that level of abstraction there is no obvious reason to say that Merge has anything to do with natural language at all. The fact that it is so successfully used in language is not a strong argument for the idea that the biological purpose of Merge is linguistic. Since Gould and Lewontin (1979), the fallacy of reading cultural functions into biological structures is known as "Panglossianism."

The crux of the matter is that Merge, in abstraction of its lexical application, is not linguistically functional. It is for a reason that before the shift to syntactocentrism, the morpheme was generally seen as the smallest linguistically functional unit. This is the wisdom behind Saussure's idea that the sign with its dual nature (sound-meaning) is the core element of language. Our combinatorial capacities (as expressed by Merge) are extremely important for language as we know it. No matter how powerful it makes language as a tool, it is only an auxiliary facility. There is just as little reason to call the capacity for Merge the "faculty of language in the narrow sense" (FLN in Hauser, Chomsky and Fitch 2002) as there is reason to call a gasoline engine "a car in the narrow sense."

7. Reconciling biology with culture

The currently biolinguistic interpretation of theories of grammar, it seems to me, is only possible due to an error that also led pre-1970 generative grammar astray: the shift from a traditional, sign-based theory of grammar to a syntax-based theory (where the notion "syntax" was borrowed from the lexicon-independent algorithms designed for formal languages). The public cultural objects stored in the lexicon are the *sine qua non* for language, no matter how narrowly conceived. Without the function assigned to them by words, the hierarchical, recursive structures of syntax have no linguistic significance. This makes the project of a grammatical meta-theory in terms of individual psychology or biology (in the non-trivial sense) futile.

I agree with Chomsky's distinction between empiristic and rationalistic theories of language acquisition and his rejecting of extreme versions of the former, like Skinner's behaviorism. Like Chomsky's own theories, such theories are selection theories but with a hypothesis space so unconstrained that language acquisition, both in its speed and in its convergence on deep universal properties of the various languages, would become impossible. Poverty-of-the-stimulus arguments strongly suggest that language exploits innate structure.

However, nothing justifies a Panglossian interpretation of the innate structures involved. In a strictly biological context, structures have become functional in the course of evolution either by adaptation or by exaptation. Adaptation is the gradual adjustment of structures to particular functions by natural selection. Exaptation is a term coined by Gould and Vrba (1982) and was called pre-adaptation in earlier theories. In fact, it is a phenomenon that was discussed by Darwin himself and discussed in the German-speaking world of the 19th century as *Funktionsverschiebung* (see Russell 1916). What it means is that a structure was first selected for one function and was subsequently applied and adjusted to another function, also by natural selection. An example given by Gould and Vrba is the wings of birds that originally evolved as thermo-regulators. I think the notion of exaptation is extremely

important because it illustrates the fact that there is no intrinsic relation between form and function. It is not predictable from physical principles which function a given structure will take in the future. Kauffman (2007) rightly takes that as an argument against reductionism, in this case the idea that biology can be reduced to physics.

How does language fit into this picture? Hauser, Chomsky and Fitch (2002) suggest that the faculty of language in the narrow sense (FLN) only includes recursion and that it may have evolved originally for reasons other than language. Possibilities they mention are number, navigation and social relations. Basically, then, they seem to claim that linguistic functionality could be a matter of exaptation.

However, it seems clear to me that neither adaptation nor exaptation will do for linguistic functionality, for the simple reason that it is based on an invention and therefore on human agency, a phenomenon found nowhere else in the biosphere. What comes to mind here is a distinction that was made by Searle (1995: 20), namely between agentive and non-agentive functionality. Examples given by Searle are the heart and a paperweight. The functionality of the heart developed by natural selection and comes about during ontogeny on the basis of our genetic program, without human interference. When we use a stone as a paperweight, however, its function is not a matter of natural selection but of human decision. We can even use an object designed for some other purpose as a paperweight, as long as it has the right size and weight. That would show an agentive version of exaptation. Agentive functionality, it seems, is the right notion for language, even in the narrowest sense. This situates language outside the scope of standard biology, which has no equivalent of agentive functionality (apart perhaps from some very rudimentary use of tools among animals).

Words are tools and tools are the prototypical examples of physical structures showing agentive functionality. Moreover, words are not just tools but *cognitive* tools, something unknown even among our closest relatives the great apes. Just as standard tools are an extension of our body, cognitive tools are an extension of our minds. Humans are not just standard biological entities but beings that live in symbiosis with objects that extend the sphere of their intentionality. Also within the human body, we can make a distinction between structures that fall within our sphere of intentionality and structures that do not. The hands, for instance, can be used as tools, while the kidneys cannot. The lungs, as discussed earlier in connection with trumpet playing, are an interesting case in that they have a non-agentive function in respiration but an agentive function in the playing of wind music. The reason is clear: unlike the heart and the kidneys, the functioning of the hands and the lungs is accessible to our control and therefore falls within the sphere of human intentionality.

These considerations also apply to the brain. Much of the processes going on in the brain is not accessible to us and therefore falls outside the sphere of human intentionality. Words and their use, however, are accessible to consciousness and willful arrangement and therefore do fall within the sphere of human intentionality. This does, of course, not mean that we have conscious access to *how* the brain manages the use of words, but that is true for all tool use, including the use of a hammer or a computer. We have no idea what happens in our brain when we use a hammer. In order to be used as a tool, it suffices that a structure is under our control in one way or another.

The most important preliminary conclusion at this point is that human agency and intentionality cannot be characterized by the standard forms of functionality known from biology (*viz*., adaptation and exaptation). The missing concept was Searle's agentive functionality. Recent developments in neurobiology give further substance to this key notion for the understanding of the biological foundations of culture. In a very important book, Dehaene (2009) gives a neurobiological account of another form of cognitive technology, our use of writing systems. Writing and reading are interesting because it is entirely uncontroversial that writing is a relatively recent invention, say, of 5000 years ago. That is far

too short a period for a capacity to have developed by natural selection. Nevertheless, as Dehaene shows, our use of graphic systems (of all cultures!) is governed by very specific areas of the brain, reminiscent of the classical areas of Broca and Wernicke. Dehaene and others have identified a region in the occipito-temporal sulcus of the left hemisphere that Dehaene has dubbed "the letterbox."

This is a very important discovery because it shows that even uncontroversially cultural phenomena, like writing and reading, are not controlled by some mechanism of general intelligence, but by very specific areas evolved for very specific purposes. In this case, the area in question is specialized for the recognition of shapes and for object invariance (i.e., the capacity to recognize something as the same object when seen from different angles and perspectives). This is largely an evolved, innate capacity, with obvious survival value for apes and monkeys as well (hence the finding of homologous structures in their brains). However, we humans have been able to give these innate structures a new function, namely the function created by the invention of writing systems. This is the agentive form of function assignment discussed earlier and Dehaene calls it our capacity for "recycling:" our capacity to give very specific, innate brain structures a new function by using them in a new, culturally invented context.

Recycling, in Dehaene's sense, is precisely the notion we are looking for and I think it sets a standard for all cognitive phenomena with a cultural, agentive component, including natural language. I see both spoken language and writing systems as cognitive technology, primarily memory technology. Derived from that, we have a set of linguistic tools for the support of thinking and for communication. If both are cognitive technologies, an interesting question arises as to the differences. Spoken language differs from writing, after all, in that it is universal and acquired much earlier in life and practically without explicit instruction. It seems to me that none of that makes a difference for the logic of the problem. Nothing in the recycling hypothesis says anything about the fact that some forms of recycling are easier to obtain than others. Singing songs, for instance, is universal and acquired early in life, while playing the piano is far from universal, learned later in life and with much more effort and explicit instruction. Nevertheless, both are uncontroversially cultural activities making use of innate capacities. It is just that some activities lean more heavily on easily accessible innate capacities than others, as can be demonstrated with numerous examples.

Another possibility is hypothesizing that language involved a certain measure of *coevolution* (Deacon 1997). That, too, is entirely compatible with the recycling hypothesis and I even consider it likely that the speed of access and use of language was favored by natural selection. However, that would be facilitating evolution, not function-creating evolution. Recall that the function of organs like the heart is created by evolution, while linguistic functionality is created by human invention. Nothing of principle excludes the possibility that the use of inventions is "rewarded" by natural selection.

8. Concluding remarks

Altogether, then, I think that Dehaene's notion of recycling is the right concept for human capacities that integrate biological and cultural phenomena. It is confirmed in case after case and is the cutting-edge idea about the relation between the brain and its cultural applications (see also Marcus 2004: 140). It avoids Panglossianism and the preformationist idea that we can meaningfully speak of a Universal Grammar (LAD) even before the child is exposed to the words that are a necessary condition for calling something "linguistic" in the first place. In general, there is no evidence for innate biological structures in the brain with intrinsic cultural dedication. Our internal organs are functionally dedicated thanks to evolution, but the brain

structures involved in language are functionally dedicated by human agency, which demonstrated its power in the invention of words shared by a community.

Preformationsm in biology was the idea that what, say, elephants were at earlier stages of their ontogeny was just smaller elephants, up to the very small ones present in their father's sperm.[3] Preformationism has since long been replaced by a more epigenetic approach to embryology: what precedes elephants in their ontogeny is not elephant-like all the way back, even if there is substantial predetermination in that the process is largely controlled by elephant DNA. This lesson from biology seems to apply to language acquisition as well: there are no good reasons to assume full-fledged concepts or a language of thought before birth. Probably, then, it is not the Panglossian and preformationist rationalism of Fodor and Chomsky that is the ultimate answer to Skinner and his successors but rather the epigenetic rationalism of Jean Piaget.[4] According to Piaget, mental growth proceeds in stages that are qualitatively different from one another. This approach better suits the fact that we are born with a rich, genetically determined mental endowment but that nothing forces us to assume a *linguistic* mental state before actual exposure to the words of our community.

[3] For a historical perspective, see Westfall (1977:100) and Hankins (1985:141).
[4] See Piattelli-Palmarini, ed. (1980) and also Koster (2009a).

Bibliography

Boeckx, C. and K. Grohmann. 2004. Putting phases into perspective. Paper presented at Tools in Linguistic Theory, Budapest, May 2004.

Brame, M. 1978. *Base generated syntax*. Seattle: Noit Amrofer.

Chomsky, N. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, N. 1959. Review of B.F. Skinner, *Verbal behavior*. *Language* 35: 26-57.

Chomsky, N. 1964. *Current issues in linguistic theory*. The Hague: Mouton.

Chomsky, N. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.

Chomsky, N. 1970. Remarks on nominalization. In R. Jacobs and P. Rosenbaum (eds.), *Readings in English transformational grammar*. Waltham, Mass.: Ginn and Company, 184-221.

Chomsky, N. 1971. Deep structure, surface structure and semantic interpretation, in D. Steinberg and L. Jakobovits (eds.), *Semantics: an interdisciplinary reader in philosophy, linguistics and psychology.* Cambridge: Cambridge University Press, 183-216.

Chomsky, N. 1973. Conditions on transformations, in S. Anderson and P. Kiparsky (eds.), *A Festschrift for Morris Halle.* New York: Holt, Rinehart and Winston, 232-286.

Chomsky, N. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, N. 1986. *Knowledge of language*. New York: Praeger.

Chomsky, N. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.

Chomsky, N. 2000. Minimalist inquiries, in R.Martin, D. Michaels and J. Uriagereka (eds). *Step by step: essays on minimalist syntax in honor of Howard Lasnik*. Cambridge, Mass: MIT Press.

Collins, C. 2002. Eliminating labels. In S. D. Epstein and T. D. Seely (eds.), *Derivation and explanation in the minimalist program*. Malden, etc.: Blackwell, 42-64.

Croft, W. 2001. *Radical construction theory: syntactic theory in typological perspective*. Oxford: Oxford University Press.

Culicover, P. and W. Wilkins. 1984 *Locality in linguistics theory*. Orlando: Academic Press.

Culicover, P. and R. Jackendoff. 2005. *Simpler syntax*. Oxford: Oxford University Press.

Deacon, T. 1997. *The symbolic species: the co-evolution of language and the brain*. New York: W.W. Norton & Company.

Dehaene, S. 2009. *Reading in the brain: the science and evolution of a human invention.* New York: Viking.

Emonds, J. 1970. Root and structure preserving transformations. PhD dissertation. Cambridge, Mass.: MIT.

Evans, N. and S. Levinson. 2009. The myth of language universals: language diversity and its importance for cognitive science. *Brain and Behavioral Sciences* 32: 429-492.

Fiengo, R. 1974. Semantic conditions on surface structure. PhD dissertation. Cambridge, Mass.: MIT.

Goldberg, A. 1995. *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Gould, S. J. and R. Lewontin. 1979. The spandrels of San Marco and the panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London*. B 205 (1161): 581–598.

Gould, S. J. and E. Vrba. 1982. Exaptation: a missing term in the science of form. *Paleobiology* 8: 4-15.

Hankins, T. 1985. *Science and the Enlightenment*. Cambridge: Cambridge University Press.

Hauser, M., N. Chomsky, and T. Fitch. 2002. The faculty of language: what is it, who has it, and how did it evolve? *Science* 298: 1569-1579.

Higgins, R. 1973. The pseudo-cleft construction in English. Cambridge, Mass.: PhD dissertation MIT [Published by Garland, New York, 1979].

Hornstein, N., J. Nunes, and K. Grohmann. 2005. *Understanding minimalism*. Cambridge: Cambridge University Press.

Jackendoff, R. 1969. Some rules of semantic interpretation for English. PhD dissertation. Cambridge, Mass.: MIT.

Jenkins, L. 2000. *Biolinguistics*. Cambridge., Mass.: MIT Press.

Kathol, A. 2000. *Linear syntax*. Oxford: Oxford University Press.

Katz, J. and P. Postal. 1964. *An integrated theory of linguistic descriptions*. Cambridge, Mass.: MIT Press.

Kauffman, S. 2007. Beyond reductionism: reinventing the sacred. *Zygon* 42: 903-914.

Kayne, R. 1984. *Connectedness and binary branching*. Dordrecht: Foris.

Kayne, R. 2005. *Movement and silence*. Oxford: Oxford University Press.

Klima, E. 1964. Negation in English. In J. Fodor and J. Katz (eds.), *The structure of language: readings in the philosophy of language*. Englewood Cliffs NJ: Prentice Hall, 246-323.

Koster, J. 1978. *Locality principles in syntax*. Dordrecht: Foris.

Koster, J. 1987. *Domains and dynasties: the radical autonomy of syntax*. Dordrecht: Foris.

Koster, J. 2007. Structure-Preservingness, internal Merge, and the strict locality of triads, in S. Karimi, V. Samiian, and W. Wilkins (eds.), *Phrasal and clausal architecture: syntactic derivation and interpretation*. Amsterdam/Philadelphia: John Benjamins, 188-205.

Koster, J. 2009a. Word meaning and the preformationist fallacy. Groningen: MS., University of Groningen. http://www.let.rug.nl/koster/papers/Word%20meaning%20and%20the%20preformationist%20fallacy%201.pdf

Koster, J. 2009b. IM not perfect: the case against copying. Groningen: Ms., University of Groningen. http://www.let.rug.nl/koster/papers/IM%20not%20perfect1.pdf

Lenneberg, E. 1967. *Biological foundations of language*. New York: John Wiley & Sons, Inc.

Maling, J. 1983. Transitive adjectives: a case of categorial reanalysis, in F. Heny and B. Richards (eds.), *Linguistic categories: auxiliaries and related puzzles*. Vol. I. Dordrecht: D. Reidel, 253-289.

Marcus, G. 2004. *The birth of the mind*. New York: Basic Books.

May, R. 1977. *The grammar of quantification*. Cambridge, Mass.: PhD dissertation MIT.

May, R. 1985. *Logical form: its structure and derivation*. Cambridge, Mass.: MIT Press.

Muysken, P. 1982. Parametrizing the notion "head." *Journal of Linguistic Research* 2: 57-75.

Piattelli-Palmarini, M. (ed.). 1980. *Language and learning: the debate between Jean Piaget and Noam Chomsky.* Cambridge, Mass.: Harvard University Press.

Pollard, C. and I. Sag. 1994. *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.

Reinhart, T. 1976. The syntactic domain of anaphora. Cambridge, Mass.: PhD dissertation MIT.

Riemsdijk, H. van. 1978. *A case study in syntactic markedness: the binding nature of prepositional phrases.* Dordrecht: Foris.

Rizzi, L. 1990. *Relativized minimality*. Cambridge, Mass.: MIT Press.

Rosenbaum, P. 1967. *The grammar of English predicate compliment constructions*. Cambridge, Mass.: MIT Press.

Ross, J. R. 1967. Constraints on variables in syntax. Cambridge, Mass.: PhD dissertation MIT.

Russell, E. 1916. *Form and function: a contribution to the history of animal morphology*. London: Murray

Saussure, F. de. 1916. *Cours de linguistique générale*. C. Bally and A. Sechehaye, (eds.). Lausanne and Paris: Payot.

Searle, J. 1995. *The construction of social reality*. London, etc.: Penguin Books.

Tesnière, L. 1959. *Éléments de syntaxe structurale*. Paris: Editions Klincksieck.

Tomasello, M. 2009. *Origins of human communication*. Cambridge, Mass.: MIT Press.

Wasow, T. 1972. Anaphoric relations in English. Cambridge, Mass.: PhD dissertation MIT.

Westfall, R. 1977. *The construction of modern science: mechanisms and mechanics*. Cambridge: Cambridge University Press.