Chapter 21

Second thoughts about the Chomskyan revolution

Jan Koster
University of Groningen

1 Introduction[1]

According to a widespread belief, the field of linguistics was revolutionized since the middle of the previous century by the ideas of Noam Chomsky at the Massachusetts Institute of Technology in the USA. This revolution, as I will discuss in this article, is a myth. To be sure, to prevent misunderstanding at the outset, it has to be said that linguistics in the intended sense has been a great success. In contrast with the often dull field of the 1950s, Chomskyan linguistics has rejuvenated the field and even given it intellectual mass appeal occasionally. The field has not only exploded quantitatively but also qualitatively. We have seen growth that, no doubt, is unique in the history of the humanities. And yet, it has been my conviction for a long time that something is wrong with the field, not only in its technical development, but particularly in the way it is interpreted at a meta-theoretical level.

Before Chomsky, European structuralism had a broadly Saussurean orientation. I will argue that, technically speaking, the first 20 years of generative grammar, far from being revolutionary, showed a gradual reinvention of the structuralist wheel. Even more confused has been the persistent meta-theoretical reinterpretation of language (in some narrow sense) as a specialized biological faculty. The technical problem and the meta-theoretical problem are largely independent but, in practice, conspire to create the current theoretical stagnation.

Before going on, it is useful that I first give a short summary of the relevant aspects of the Saussurean heritage, contrasting it along the way with some central Chomskyan tenets. The reader should keep in mind that I am mostly focused on syntax

[1] I would like to thank Christina Gagné for discussion and helpful comments on an earlier draft of this article. All remaining errors are my own.

and semantics and that I leave the evaluation of developments in phonology happily to others. According to Saussure, language (in he narrowest possible sense) is not an individual-psychological or biological faculty (as for Chomsky) but a system of signs. Signs are invented artifacts and, as such, belong to the shared culture of some language community (see also Popper 1972).

Signs according to Saussure have three dimensions, the best known of which are the overt *signifiant* (audible visible, readable, whatever is the sensory mode) and the covert *signifié* (some conceptual substance). Trivially, the relation between *signifiant* and *signifié* is conventional, as has been a common idea at least since Aristotle.

The second covert dimension of the Saussurean linguistic sign is often overlooked and is about the relations between a simple sign and its environment. Thus, every competent speaker who knows the English word *book* also knows that it can be preceded by an article: *the book*. Part of a linguistic sign, then, can be "under water". Thus, we have signs like *Art book*, *book PP*, *Art book PP*, etc. (where variable *Art* stands for article and variable *PP* for prepositional phrase). Each sign has a set of more or less fixed, invisible environments, which can be made visible by substituting the variables by constants: *the book*, *books from Geneva*, *the books from Geneva*, etc. The predictable environments of signs fall under what Saussure called "syntagmatic relations" and were later called the "valency" of a sign (mostly applied to verbs, see Tesnière 1959). Syntactic structures, then, are complex signs, which spell out the environmental properties of simpler signs. Both simple signs and complex signs conform to the conventions of the "trésor commun", the non-individual "langue."

Needless to say, our conventions are constrained by our individual biological properties, for the same trivial reasons as why all our forms of culture are constrained by our nature. Note also, that there are no specific derivational levels that can be referred to as "the interfaces". Each linguistic sign, including the simplest morpheme, is a three-fold interface, connecting "signifiant", "signifié" and conventional combinatorial potential ("syntagmatic relations"). There are also sign systems, like traffic signs, that miss a significant syntagmatic dimension and are therefore only interfacing two dimensions (*signifiant* and *signifié*).

2 Some theories of generative syntax since the 1950s

2.1 (Extended) Standard Theories

The following three are the most important versions of Chomskyan generative grammar since the 1950s:

(1) a. (Extended) Standard Theories (1955–1973)
b. Government–Binding Theories (1973–1995)
c. Minimalist Theories (1995–)

This is a rough periodization which cannot be exact for the simple reason that key ideas from the various periods overlap. The idea of a Chomskyan revolution is particularly based on the much hyped first period (1a). It started in 1955 with the

voluminous *The Logical Structure of Linguistic Theory*, was popularized by *Syntactic Structures* (1957) and culminated in *Aspects of the Theory of Syntax* (1965), the so-called Standard Theory. This period was concluded by the Extended Standard Theory of Chomsky (1971) and the lexicalism of Chomsky (1970).

The next period (1b) sought to improve explanatory adequacy further by limiting the initial hypothesis space not only by X-bar theory (Chomsky 1970) but, most typically, also by "conditions on transformations", as in the paradigmatic articles Chomsky (1973) and Chomsky (1977). Movement rules, the most important transformations in this framework, were gradually reduced to the schema "Move Alpha". This type of theorizing culminated in another paradigmatic work, *Lectures on Government and Binding* (1981) (GB). In spite of the fact that GB-style theorizing was gradually hoped to be superseded, particularly after Chomsky (1995), by minimalist speculations (1c), "normal science" in generative linguistics is largely determined by GB-type analyses, up until the present day. Many linguists are only terminologically affected by Minimalism, replacing "Move Alpha" by "internal Merge" and re-baptizing good old bounding domains as "phases" (yes, I know, there are subtle differences).

Many generative linguists believe that (1a-c) is a continuous story of progress, which started with the glorious revolution of the late 1950s. In my opinion, this self-image of the field is false. Only the first period (1a) saw some potentially revolutionary ideas, but by the early 1970s it was, or should have been, clear that they were all ill-conceived. For concreteness' sake, consider the following tenets of the early theories:

(2) a. syntax-based theories instead of sign-based theories
b. two-step sentence generation: PS-rules (kernel) and transformations
c. multiple levels of representation: deep and surface structure
d. use of formal methods (mathematics, logic)

About (2d) we can be brief: after some initial results, mathematical linguistics practically disappeared from Chomskyan practice since the 1970s.This is ironical because the initial interest in mathematics was a formidable tool of propaganda in establishing the field's prestige and image of revolutionary paradigm shift. Since the 1970s, interest in mathematics and logic mostly lived on in formal semantics (Generative Semantics, Montague-inspired work, etc.).

From a Saussurean structuralist's point of view, talking about syntax independent of the properties of signs is bound to be a failure. And so it happened. Syntactic Structures (1957) generated structures without looking at the internal properties of lexical items. However, as soon as a lexicon is added, it appears that the PS-rules mimic the valencies that are also spelled out as internal properties of the lexical items (Chomsky 1965). Chomsky (1970), and particularly Chomsky (1981) drew the obvious conclusion, namely that doing things twice can only be prevented by projecting syntactic structures directly from lexical items. This was no revolutionary new insight but the reinventions of the Saussurean wheel, according to which syntactic structures are defined as the syntagmatic properties (valencies) of signs. Exit revolutionary tenet (2a).

With (2a) and (2d) gone, the idea of a Chomskyan revolution had little to boast beyond the interrelated (2b) and (2c). These were perhaps the most characteristic features of the intended revolution. Especially the term "deep structure" was of great propagandistic value at the time, obscuring the fact that it stood for something close to and as unexciting as the "strings underlying kernel sentences" in Chomsky (1957).

The history of (2b) and (2c) is very interesting. The prima facie revolutionary innovation of generative grammar was in the transformations, not in the PS-rules. The transformations were tightly connected with the new idea of multi-level representation, as they formed the core of the mapping from deep structure to surface structure. Given the fact that these were key concepts of the alleged revolution, it is, in retrospect, astonishing how fast transformations disappeared from the theoretical scene around 1970. Major classes of transformations, like pronominalizations and equi-NP-deletion, appeared to be ill-conceived and were replaced by rules of construal, the essence of which is the local sharing of properties (see Koster 1987: 8ff). Before long, construal was demonstrated to be also the better alternative for movement transformations (with NP-movements and Wh-movements as the major families; see Koster 1978).

The end of transformationalism was considerably speeded up by insights about structure-preservingness. Emonds (1970, followed by some later elaborations) demonstrated that major movement transformations were structure-preserving in the sense that they created structures that could be independently created by the bases rules (X-bar schemata). Others, including Chomsky, argued that movement rules left a "trace", like the empty NP-object in (3):

(3) *What* did you see [$_{NP}$ ...]

Of course, this very terminology of structure-preservingness and "traces" is odd in retrospect, as what we see is just "base-generated" *What* (for instance, as the Spec of CP) and the "trace" [$_{NP}$ ...] as the spelled out valency of *see*. The then current terminology presupposed transformationalism, while what actually was demonstrated was that transformationalism is false, also for "movement" rules. Altogether, then, it is fair to say that transformationalism was as dead as a doornail by the mid 1970s.

This outcome should have had serious consequences for how we evaluate the so-called Chomskyan revolution. The least we can say is that, in retrospect, it was more hype than substance. However, I think we should go one step further: it created perhaps impressive enthusiasm and empirical momentum but conceptually, it was (at least until the mid-1970s) a complete failure and left the field more or less where it was before the mid-1950s. With transformations gone (2b), multiple-level representation (2c) was gone as well. What remained was X-bar structures projected from the lexicon as the foundation for property sharing (construal). This is just a variant of traditional *Wortgruppenlehre*, entirely compatible with Saussurean and other pre-1950s structuralist assumptions (see also Ries 1928, De Groot 1949 and, with a more English-oriented account of the parts of speech, Jespersen 1924).

2.2 Government–Binding and Minimalist Theories

With the generative enterprise diminished to a more or less traditional *Wortgruppenlehre*, the focus of the field shifted to "principles and parameters", with the ambition to restrict the class of possible grammars to a very few or even one, with open parameters to account for the differences among languages. The theory of parameters never was developed beyond superficial descriptions close to the data, like the VO/OV-parameter supposed to account for the difference between VO- and OV-languages. Principles were often empirical generalizations for specific domains, like bounding theory for "movements" and binding theory for anaphoric constructions. The "principles and parameters" framework has had two varieties so far, the Government-Binding versions (1b) and the Minimalist versions (1c). For present purposes, it is important to keep in mind that the GB-versions are based on X-bar theory and therefore a continuation of the traditional *Wortgruppenlehre*. This was implicitly denied by Chomsky, as will be clarified in a moment. Minimalist theories are more radical in that their core is a step away from traditional word groups and constructions. This core is the Galilean grail known as "Merge", which produces word groups only through interactions known as "mappings to the interfaces" (sensori-motoric and conceptual-intentional).

The most important question from the current perspective is whether the failed revolution of the first period (1a) got a new chance in the "principles and parameters" framework. The short answer is: no, the field gradually disintegrated into conceptual chaos. There were plenty of linguists who took the consequences of the developments of around 1970 and got rid of transformationalism altogether, thereby implicitly recognizing the failure of the revolution (Brame 1978; Bresnan 2001; Pollard & Sag 1994; Koster 1978; 1987). Chomsky, in contrast, preferred denial and sought to save transformationalism with the transformational residue "Move Alpha". This enabled him to leave the revolutionary illusion intact, with a full-fledged defense of multiple levels of representation (from then on called D-structure, S-structure and Logical Form (LF); see Chomsky 1981 and Koster 1987 for a critique). Never mind that these levels later on disappeared through the minimalist backdoor, generously ignoring the fact that the relevant insight had been available for decades.

The most representative sub-theory of the GB-framework is the so-called bounding theory, with Subjacency as its core principle. So, if we want to know if generative grammar ever deserved the predicate "revolutionary" after all, it is useful to focus on Subjacency. We would minimally expect Subjacency to be non-trivial and innovative. It appears that Subjacency, and the problems it seeks to solve, is an artifact of the transformational framework that was already about to collapse by the time Subjacency was introduced in Chomsky (1973).

According to Subjacency, no transformational rule involves two categories X an Y across **two** bounding nodes. The bounding nodes were NP and S' (or the parametric variant S), which were later re-baptized as DP and CP (IP). Subjacency was supposed to be a more principled, deeper replacement of the famous, but somewhat disparate list of island constraints of Ross (1967), which in turn replaced the more appealing but empirically inadequate A-over-A Principle of Chomsky (1964).

Two questions immediately come to mind with respect to Subjacency in the current context: 1) is it empirically adequate, and 2) does it exceed the conceptual boundaries of traditional, pre-Chomskyan *Wortgruppenlehre*. I do not have the space here to do full justice to these questions, but in summary, answers amount to the following. As for empirical adequacy, Subjacency appeared to be wrong on two counts. First of all, the number two (highlighted in the description of Subjacency above) appeared to be unmotivated. Whenever Subjacency applies in the relevant contexts, one node is enough. Thus, the relations anachronistically described by movement transformations, in the unmarked case, never apply across single NP (DP)-boundaries, as correctly stipulated by the NP-constraint of Bach & Horn (1976). Never in such cases, a second boundary node (like S' (CP), S (IP) or a second NP (DP)) has to be specified. Second, an interesting, largely ignored discovery was made as soon as other languages were taken into account. Thus, largely on the basis of Dutch but with an eye for lots of other languages, van Riemsdijk (1978) concluded that PP had to be added to the inventory of bounding nodes. Koster (1978) generalized bounding to all lexical categories (maximal projections, extended XPs in the sense of Grimshaw 2000).

The significance of these discoveries has been completely misunderstood. It was implicitly shown that, far from being the effect of a fancy 2-node condition on fancy transformational rules, bounding phenomena showed an effect that could have been obvious from traditional *Wortgruppenlehre*, namely that the "size" of a word group is determined by the valency of its lexical head. For instance, the DP complement of a P is only interpreted as such within a PP, which is the exclusive domain of the P's syntagmatic relations. In the unmarked case, then, bounding is what is to be expected on lexicalist assumptions: it is limited to maximal projections (word groups). In some languages, like English, in a very narrow range of contexts, domains can be extended further by absorption of V-, A-, or P-projections into a more encompassing V-projection (see Koster 1987).

In short, the core principle of the next "revolutionary" stage, GB-theory, is reducible to what follows from the traditional definition of a word group.

What about Minimalism (1c, Chomsky 1995)? The core notion of Minimalism is recursive Merge (plus mappings to the interfaces). The potential revolutionary character of Merge is believed to rest on its radical "Galilean" nature, i.e., its status of a perfect object that only produces the "messy" data in interaction with external factors. In principle, this seems to be an interesting move, but in practice it comes down to the observation of properties that are hardly controversial and entirely compatible with most theories of syntactic structure, both traditional (like *Wortgruppenlehre*) and modern (like Construction Grammar).

First of all, it should be noticed that Merge is like pre-lexicalist, pre-1970 grammar in that it introduces hierarchical recursive structure independent of the lexicon. But since the merged objects also have hierarchical, recursive structures as their valency, Merge reintroduces the redundancy that X-bar theory was designed to get rid of. This does not look like progress. However, one can maintain that Merge at least introduces the right properties: binary branching and adjacency of the related cate-

gories. Reformulated as a constraint on representations, the substance of Merge is preserved and the redundancy problem is solved. As before, syntactic structures are projected from the lexicon (instead of being generated with Merge) and they are accepted as long as they conform to the constraints formerly derived from Merge. I have shown elsewhere that this view of Merge as a meta-constraint on representations has other interesting consequences as well, as stated in the theory of triads (see Koster 2007; 2015).

For the belligerent among us, this is bad news because it makes the properties of Merge compatible with almost any framework on the market. Theorists of Construction Grammar, for instance, can postulate constructions without fear, as long as their constructions conform to the locality constraints formerly contributed to Merge.

Most significantly, the locality constraints are also entirely compatible with pre-revolutionary *Wortgruppenlehre*. Again, there is no reason to complain about the descriptive-analytic productivity of the field since the mid 20th-century, but the idea of revolutionary conceptual innovation is in urgent need of revision. The revolution was mainly a matter of hype and rhetoric. As for substance, there is almost complete continuity with the tradition. Philosophically speaking, the pre-WWII European ideas about language were even significantly better, as will be discussed next.

3 Against biolinguistics

The main shift since the 1950s has been from an external, socio-cultural concept of language to language seen as a matter of an individual faculty, also referred to as a genetically determined universal grammar (UG) or as I-language (where I stands for individual, internal and intensional). Ultimately, it is hoped, I-language can be unified with biology, whence the term "biolinguistics" (see Lenneberg 1967; Jenkins 2000; Hauser et al. 2002).

Inspired by Saussurean ideas, the focus of pre-Chomskyan linguistics in Europe was on the collective "langue", a public system of signs serving the needs of symbolization. Of course, individual aspects were also recognized, but those were relegated to "parole" and seen as outside the language system in the narrow sense. Seeing language as a phenomenon external to the individual was the common view. Major philosophers, like Karl Popper, for instance, saw language as a "World 3" phenomenon (see Popper 1972). All of this was in accordance with the insight that one of the most characteristic aspects of human minds is their living in symbiosis with shared external memories (see Donald 1991). Language was somehow seen as the pivotal phenomenon bringing this symbiosis about. I think this view is correct and that the view of language as an individual faculty is wrong.

Language is a socio-cultural phenomenon because it is based on morphemes and words. Words do not grow from trees or in wombs but are artifacts invented by someone and adopted by the community, usually of the inventor (with exceptions such as loanwords). Words thus invented and adopted are maintained as part of the community's cultural record, in the form of oral or written traditions. Morphemes and words have, as we saw, a valency that can be lexicalized to form more complex

signs. These valency patterns are also part of the record of a language community, even if many community members are not aware of that. However, if somebody, on purpose or by accident, deviates from the accepted norm, most community members will notice that. Valency patterns are recursive because each realized pattern opens up new slots with new realizable contexts, etc., *ad infinitum*.

Needless to say, such complex cultural objects with recursion, can only be handled by brains that are able to do so. However, from the relation between the ability and complex objects, it does not follow that the ability in question is a language ability. The relation is not intrinsic but accidental, a relation also known as an application. Applications must have an agentive cause, a context that somehow give the object a function. The distinction that we commonly make between biological and cultural applications (or functions) depends on the nature of the agentive cause. If the cause is a quasi-agent, like natural selection, we call the application 'biological'. Examples of biological functionality are the natural functions of the organs of the body. Similarly, so-called bio-computation, like the kind involved in mammalian vision, is biological because the functionality is not caused by human intervention.

However, if the agentive cause is human invention, we commonly call the application 'cultural'. Interesting examples are those that involve both agentive- and non-agentive functionality. Consider organs like the lungs. The lungs have an obvious biological function brought about by non-agentive causation, for instance, by natural selection. However, if we wish we can also give the lungs a cultural application, mediated by human-made artifacts, like trumpets and other wind instruments. This example shows that the status of the application (biological or cultural) has nothing to do with the innateness of the structures involved. Obviously, the lungs are innate in the relevant sense. Nor does it matter how many applications there are. In this case, the lungs have exactly one cultural application.

The only thing that matters is the nature of the agentive cause of the application. By that criterion, Chomsky has it wrong when he compares linguistic functionality with the functionality of the organs or the computations involved in the mammalian visual system. The linguistic application of our capacity for recursion (innate or not) is agentive because mediated by human-made artifacts like the trumpet. In this case by linguistic signs, like morphemes, words or phrases.

It must be concluded therefore that biolinguistics is an untenable proposition. All things cultural, from sports to music and language, exploit the more or less innate capacities of our body and brain, crucially as a matter of free agentive application. This in contradistinction to the biological functions of the body, which do not enjoy such freedoms of application.

I will end with an argument against biolinguistics based on intentionality. Inspired by medieval examples, Franz Brentano (1874) developed the valuable insight that there is an essential difference between merely physical states and mental states. The difference is intentionality or "aboutness". Thus, an arbitrary object, like a stone is not about anything. However, the *word* "stone" involves mental states and is therefore about something, for instance about stones.

The relevant notion of aboutness deserves some further clarification. One might

say, for instance, that instruments are about something. Thus, a thermostat can be said to be about temperature and an alarm clock about time. Obviously, however, thermostats and alarm clocks do not have mental states. Insofar as these instrumental objects are about something, it is thanks to the user. Therefore, we say in these cases that the objects have "derived intentionality". In general, intentionality involves mental states that relate a target (what the states are about) with a source. The most interesting aspect of intentionality, highly relevant for cognitive science, is the question what is the source of intentionality.

Crucial is the insight that brains and brain states are tools serving the user and therefore intentional objects with derived intentionality. The question, then, is what is the source of their intentionality. There is a misguided tendency to look for the source at the brain itself ("we are our brains"). Bennett & Hacker (2003) have called this "the mereological fallacy": confusing the part (the brain) with the whole (the person using the brain). In short, people, not their instruments such as computers and brains, are the sources of intentionality.

The crux of the argument is that people (living agents) cannot be defined in purely physical or biological terms because they have, next to a physical identity a sociocultural identity and a history. The latter parts of our identity are at least as important as the former. In general, therefore, the intentionality of the mental is a decisive barrier against its naturalization in our theories of cognition.

This is directly relevant as an argument against biolinguistics. According to Chomsky, "knowledge of language" (I-language) is a state of the mind/brain. Language learning, in his view, is a development from an initial state S_0 to a relatively stable state S_s, where S_0 = UG (see, for instance, Chomsky 2007). However, it is a serious error to say that knowledge of language is a brain state. This would be of the same calibre as saying that knowledge of time is a state of your alarm clock. Like clock states, brain states have derived intentionality and only represent knowledge thanks to the necessary source and target of their intentionality. We have already seen that the source of human intentionality cannot be naturalized. This suffices to make "biolinguistics" an unattainable goal. However, things are much worse for biolinguistics. As a state of the mind/brain S_s, Chomsky's "knowledge of language" is intentional, therefore not only has a source but also a target (what the mental state is about). What, then, is Chomsky's state S_s actually about? It never has become clear.

One thing is certain, however: there is no known biological answer to the question what language-as-a-mental state is about. Luckily, as was proposed earlier on in this article, there is a perfectly satisfying, traditional structuralist answer: knowledge of language is about systems of complex signs (Saussure: 'langue') and their use ('parole'). If this answer is correct, biolinguistics must also be rejected from the target side of the intentionality relation. This is so, I can only repeat, because signs are not growing from trees or in wombs but are inventions by human agents and preserved in our cultural records.

4 Conclusion

The history of Chomskyan generative grammar is in urgent need of revision. All what seemed revolutionary about it in the 1950s and 1960s turned out to be untenable, often as early as in the 1970s. Later revisions failed to reanimate the revolution and were more than once a step back in the direction of pre-Chomskyan models of grammar, leaving a GB-style analytic-descriptive kind of normal science, with theoretical notions mostly compatible with both traditional and more recent kinds of frameworks. The "biolinguistic" effort of recent years is doomed to failure, as it continues the fundamental error of seeing language (in some narrow sense) as an individual mental state rather than as a Sassurean "trésor commun".

References

Bach, Emmon & George Horn. 1976. Remarks on 'Conditions on transformations'. *Linguistic Inquiry* 7. 265–299.

Bennett, Maxwell & Peter Hacker. 2003. *Philosophical foundations of neuroscience.* Malden, MA: Blackwell.

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

Brentano, Franz. 1874. *Psychologie vom empirischen Standpunkte.* Leipzig: Von Duncker & Humblot.

Bresnan, Joan. 2001. *Lexical-functional syntax.* Malden, Mass.: Blackwell.

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

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

Chomsky, Noam. 1965. *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chomsky, Noam. 1970. Remarks on nominalization. In Roderick Jakobs & Peter Rosenbaum (eds.), *Readings in English transformational grammar*, 184–221. Waltham, MA: Ginn & Company.

Chomsky, Noam. 1971. Deep structure, surface structure and semantic interpretation. In Danny Steinberg & Leon Jakobovits (eds.), *Semantics: an interdisciplinary reader in philosophy, linguistics and psychology*, 183–216. Cambridge: Cambridge University Press.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen Anderson & Paul Kiparsky (eds.), *A festschrift for Morris Halle*, 232–286. New York, etc.: Holt, Rinehart & Winston.

Chomsky, Noam. 1975. *The logical structure of linguistic theory.* New York: Plenum Press.

Chomsky, Noam. 1977. On Wh-movement. In Peter Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal syntax*, 71–132. New York: Academic Press.

Chomsky, Noam. 1981. *Lectures on government and binding: the Pisa lectures.* Dordrecht: Foris.

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

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + recursion = language?*, 1–29. Berlin: Mouton de Gruyter.

de Groot, Albert W. 1949. *Structurele syntaxis*. Den Haag: Servire.
de Saussure, Ferdinand. 1916. *Cours de linguistique générale*. Charles Bally & Albert Sechehaye (eds.). Lausanne & Paris: Payot.
Donald, Merlin. 1991. *Origins of the modern mind*. Cambridge, Mass.: Harvard University Press.
Emonds, Joseph. 1970. *Root and structure preserving transformations*. Cambridge, Mass.: MIT PhD thesis.
Grimshaw, Jane. 2000. Locality and extended projections. In Peter Coopmans, Martin Everaert & Jane Grimshaw (eds.), *Lexical specification and lexical insertion*, 115–133. Amsterdam: John Benjamins.
Hauser, Marc, Noam Chomsky & Tecumseh Fitch. 2002. The faculty of language: what is it, who has it, and how did it evolve? *Science* 298. 1569–1579.
Jenkins, Lyle. 2000. *Biolinguistics*. Cambridge, Mass.: MIT Press.
Jespersen, Otto. 1924. *The philosophy of grammar*. London: Allen & Unwin; New York: Holt.
Koster, Jan. 1978. *Locality principles in syntax*. Dordrecht: Foris.
Koster, Jan. 1987. *Domains and dynasties*. Dordrecht: Foris.
Koster, Jan. 2007. Structure-preservingness, internal merge, and the strict locality of triads. In Simin Karimi, Vida Samiian & Wendy Wilkins (eds.), *Phrasal and clausal architecture: syntactic derivation and interpretation*, 188–205. Amsterdam-Philadelphia: John Benjamins.
Koster, Jan. 2015. Aspects and beyond: the case for generalized construal. In Ángel J. Gallego & Dennis Ott (eds.), *50 years later: reflections on Chomsky's Aspects* (MIT Working Papers in Linguistics 77), 159–169. Cambridge, MA.
Lenneberg, Eric. 1967. *Biological foundations of language*. New York: John Wiley & Sons, Inc.
Pollard, Carl & Ivan Sag. 1994. *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.
Popper, Karl. 1972. Epistemology without a knowing subject. In Karl Popper (ed.), *Objective knowledge: an evolutionary approach*, 106–152. Oxford: Oxford University Press.
Ries, John. 1928. *Zur Wortgruppenlehre: mit Proben aus einer ausführlichen Wortgruppenlehre der deutschen Sprache der Gegenwart*. Prag: Taussig & Taussig.
Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, MA: MIT PhD thesis.
Tesnière, Lucien. 1959. *Éléments de syntaxe structurale*. Paris: Editions Klincksieck.
van Riemsdijk, Henk. 1978. *A case study in syntactic markedness: the binding nature of prepositional phrases*. Dordrecht: Foris.