## Review article / Rapport critique
### Forschungsbericht

## Concerning the Roots of Transformational Generative Grammar*

Pieter Seuren

Max Planck Institute for Psycholinguistics, Nijmegen

1. *Introductory remarks*

Transformational generative grammar (henceforth: TGG) as it has been known for decades now, has without a doubt found a permanent place in the history of linguistics. Despite the sometimes tortuous and not always transparent ways along which it has manifested itself over the past sixty years and the many dead alleys it got caught in, it seems most unlikely that it will disappear without a trace in the future developments of linguistic theory and practice. It is, therefore, most appropriate that ample attention should be paid to the origins of TGG as a phenomenon in 20th-century scientific thought. In this context, the recent book by Marcus Tomalin (Tomalin 2006) on the origins of Transformational Generative Grammar should kindle the interest of a wide public.

Tomalin's book *Linguistics and the Formal Sciences* (henceforth *LFS*) probes the intellectual origins and backgrounds of TGG as it appeared on the market during the late 1950s. This is an ambitious enterprise, given the enormously complex fabric of developments in philosophy (in particular ontology, epistemology and the philosophy of science), psychology, (structuralist) linguistics, logic and mathematics and their foundations during the first half of the twentieth century, and the roots of all these stormy developments in earlier centuries. In fact, the intellectual climate in which first structuralist linguistics and then its formalist successors, both in Europe and in North America, came about shows a richness, a dynamism and an intellectual power that are hard to fathom in our day. In this complex, heady jumble, Tomalin follows a particular strand, which he identifies as the roots of formal grammar, in particular of TGG.

---

Before presenting my, on the whole rather critical, view of this book, let me emphasize that the author ought to be praised for his courage and his industry in writing it. The topic is not popular, probably because the topic matter is unusually hard to get an adequate grasp of. There is thus a real need for studies of this nature, especially because the history of linguistics as a whole has been receiving much increased attention over the past few decades. *LFS* aims at filling at least one gap in our knowledge in this respect. It does so not only by calling attention to, for example, Leonard Bloomfield's (1887–1949) on the whole little known publications after 1933 or Noam Chomsky's (b. 1928) equally little known pre-1957 articles, but also by highlighting the part played by logic and mathematics in the coming about of TGG.

2.   *Survey of* Linguistics and the Formal Sciences

*Linguistics and the Formal Sciences* consists of six chapters, the first being an Introduction and the last a Conclusion. The Introduction sets the stage and the tone. The central object of interest is TGG and the central figure is Noam Chomsky. The reason given for this selection is that "during the 1970s […] it became increasingly clear that Chomsky's place in the history of linguistics was secure" (p. 7). There is, accordingly, little doubt in *LFS* regarding the brilliance and the permanent value of Chomskyan TGG in the light of Western history as a whole. Tomalin then reviews the literature on the origins of TGG (curiously leaving out Seuren 1998, which deals extensively with the subject) and concludes that it has left significant gaps — which is why he undertook to write the book.

Chapter 2 is about developments in logic and mathematics. Tomalin distinguishes three strands: *Logicism* (the Russellian view that logic forms the basis of mathematics), *Formalism* (David Hilbert's [(1862–1943)] view that the whole of mathematics can be caught in terms of a consistent axiomatic system based on arithmetic, whose expressions are formulated in a uniform formal language), and *Intuitionism* (Luitzen Brouwer's [(1881–1966)] view that all mathematics is *a priori* and thus constructible through mere thinking — and not really transmittable from one person to another). The chapter ends with a discussion of some student textbooks of the period. The emphasis throughout the chapter is on axiomatization and formalization. The point of this chapter is to sketch in broad outline the main features of those developments and debates in mathematics and logic that Chomsky could draw upon when he devised his theory of TGG.

The third, and by far the longest (53 pages), chapter is intended to home in on those "techniques or theories that were ultimately to be involved in the creation of TGG" (p. 54). For that reason it is called "Mathematical Linguistics", even though that term has come to be used for something quite different. In trying to focus on TGG, however, Tomalin takes a very broad view. Not only does he discuss the

mathematical and/or axiomatic leanings of linguists such as Bloomfield, Bernard Bloch (1907–1965), Zellig S. Harris (1909–1992), Charles F. Hockett (1916–2000), Louis Hjelmslev (1899–1965) and, somewhat surprisingly, the Tasmanian English scholar F. William Harwood (dates unknown), who is only known for a sketchy and programmatic article of four pages in *Language* of 1955 without any follow-up by Harwood himself or by others; there are also sections on recursive definitions, logical systems, constructional system theory, constructive nominalism and formal linguistic theory.

Chapter 4, though called "Systems of Syntax: 1951–1955", in fact deals only with the development of Chomsky's ideas about (the theory of) syntax during the period concerned. Next to some biography, it deals with Chomsky's 1951 MA thesis on the *Morphophonemics of Modern Hebrew* and an article he published before 1955, and also deals extensively with his 1955 bulky manuscript "The Logical Structure of Linguistic Theory", published in a heavily edited version only in 1975. The chapter also introduces the novel notion of 'constructive nominalist syntax', which, however, remains undefined and appears to denote Tomalin's interpretation of early Chomskyan TGG rather than TGG itself.

Chapter 5 is a continuation of Chapter 4, but now for the period 1955–1957. It discusses two articles published by Chomsky during this period, his reasons for rejecting stochastic grammars, and his conversion from discovery to evaluation procedures. It gives an imprecise account of how Chomsky's notion of syntactic transformation developed and reverts again to the obviously cherished issues of formalization and recursion. The concluding sixth chapter summarises what Tomalin sees himself as having achieved in the book: "… perhaps the main achievement of this book has been to associate TGG with both Formalism and Logicism, two intellectual movements that profoundly influenced scientific methodology during the first half of the twentieth century" (p. 186). It also looks ahead at the recent development of Minimalism in syntax as it has been propagated by Chomsky and some of his followers since 1995.

## 3.   *Balance and proportions*

There are a number of problems with this book, but the main and overarching problem seems to be one of perspective, balance and proportions. At least two thirds of the book are devoted to early 20th-century developments in the foundational study of logic and mathematics and the reader is made to think that that is where Chomsky took his leading ideas from. Although there is some truth in this, it is only a very partial truth. In fact, Chomsky grew up in an intellectual environment that was, in a broad and general sense, deeply influenced by positivism, structuralism, behaviorism and questions of scientific methodology, besides the much narrower sphere of the foundations debate in logic and mathematics. And

although the logico-mathematical background is essential for a proper under-standing of early TGG, it is equally essential to know about the wider background. Yet there is next to nothing in *LFS* on the philosophical questions associated with positivism and philosophy of science, which are not mentioned even once in the whole book, or on the birth of American structuralism and behaviorism. There is no mention at all of Chomsky's early dependence on Karl Popper's (1902–1994) analysis of scientific method and on his gradual (but unacknowledged) shift from Popper, via Thomas Kuhn (1922–1996) and Imre Lakatos (1922–1974), to Paul Feyerabend (1924–1994).

Moreover, in attempting to trace the logico-mathematical origins of early TGG, Tomalin has allowed himself to be led by the references and accounts pro-vided by Chomsky himself in his various writings of the period and later. In view of the many warnings in the literature that Chomsky is not to be trusted as regards the acknowledgment of his sources, one must fear that Tomalin has been a little too gullible in this respect. It is a recurrent pattern in Chomsky's writings that the authors to whom he owes most are referred to least, and then in a way that fails to reveal the true extent of his indebtedness. We shall see a few examples of this tendency below.

4.   *The algorithmic character of a generative grammar*

Tomalin's exclusive focus on the logico-mathematical aspects as presented by Chomsky himself has made for a serious distortion of historical reality. Let me start with Tomalin's failure to see the immediate link between the notion of a gen-erative grammar and Emil Post's (1897–1954) theory of algorithms, best known from his 1944 article and accessibly summarized in Chapter 4 of Paul C. Rosen-bloom's (b. 1907) *Elements of Mathematical Logic* of 1950. It should be known that, during the first half of the twentieth century, intensive debates took place, in Europe and America, around the formal foundations of, and mutual relations between, logic and mathematics — one generally speaks of the *foundations debate*. That such a debate should take place was to be expected, given the 'mathematical turn' that logic had taken since the fundamental work on the theory of algebra and its application to set theory by George Boole (1815–1864) and the subsequent work on the foundations of logic and mathematics by Gottlob Frege (1848–1925) and Bertrand Russell (1872–1970).

It requires no more than a superficial familiarity with the foundations debate to see that Chomsky's notion of a grammar as a 'machine' or 'device' generating a (possibly infinite) set $\Sigma$ of finite strings of symbols, thereby characterizing $\Sigma$, is an exact parallel to Post's notion of what constitutes an algorithm. In an attempt to formalize the general notion of an axiomatically organised theory, especially in mathematics, Post generalized the traditional notion of algorithm from the four

classic arithmetical algorithms of addition, subtraction, multiplication and division (as defined by the ninth-century Beirut mathematician Al Huárizmi, eponymus of the word *algorithm*, originally *algorism*) to that of a finite mechanical procedure or 'device' generating a (possibly infinite) set of finite strings of symbols from a given finite alphabet according to a finite set of production rules. (In fact, the term *generate* in the TGG sense originates with Post.) An algorithm **A** thus characterizes a (possibly infinite) set Σ of finite strings of symbols, whereby each string is assigned a constituent structure as a result of the application of the production rules. **A** may combine two or more outputs into a pair or *n*-tuple, upon which further operations can be carried out (as when one number is subtracted from another or two numbers are multiplied — the analog of the generalized transformations of early TGG). The set Σ may be finite or infinite. It is infinite just in case there is optional recursion in the production or transformation rules. Obligatory recursion would lead to infinite strings.

Post called a set Σ that can be generated, and thus characterized, by an algorithm **A** a *canonical* set. If there is also an algorithm **A'** that generates the complement of Σ in the alphabet given (that is, all those strings that are not generated by **A**), Σ is called *decidable*, in the sense that there is a finite algorithmic procedure for determining whether any string *f* in the alphabet given is or is not a member of Σ. Any canonical set Σ may be called a 'language' in so far as a 'language' is considered to be a well-defined set of strings of symbols. Thus, the infinite set N of natural numbers forms a 'language' on the alphabet {0,1,2,3,4,5,6,7,8,9}. This 'language' is not only canonical (as it is generated by a very simple generative algorithm, its 'grammar') but also decidable (as no combination of the alphabet's elements is 'ungrammatical' — that is, not a natural number).

It is easy to see that a TGG as conceived by Chomsky is, by definition, a Post-type algorithm applied to, or interpreted upon, linguistic forms. One recalls the first page of Chapter 2 of Chomsky's 1957 *Syntactic Structures* (p. 13):

> From now on I will consider a *language* to be a set (finite or infinite) of sentences, each finite in length and constructed out of a finite set of elements. [...] The fundamental aim in the linguistic analysis of a language L is to separate the *grammatical* sequences which are the sentences of L from the *ungrammatical* sequences which are not sentences of L and to study the structure of the grammatical sequences. The grammar of L will thus be a device that generates all of the grammatical sentences of L and none of the ungrammatical ones.

Unfortunately, Chomsky has consistently failed to point out to his readers that he owes his notion of TGG to Post (the only reference to Post I am aware of is in the hardly accessible Chomsky 1959a, in a footnote that attributes the term *generate* to Post). All Tomalin has to say about this is (p. 169):

> In particular, Post's theory of recursively enumerable sets, described in Section 3.3, provided a useful framework that could be adapted for the purposes of theoretical syntax.

The dependency of Chomskyan TGG on Post's theory of algorithms should be clear to anyone at all informed about the foundations debate of the first half of the 20th century.[1] It certainly was to the Amsterdam logician-philosopher Evert Beth (1908–1964), who wrote, having first explained the essence of Post's theory of algorithms (Beth 1963:20; translation mine: PAMS):

> I hope to be able to show that the work done by Chomsky and his associates is understood more adequately and judged more justly if one takes its background into account.

Beth then proceeds to show that the *traditional* notion of 'grammar of a language L' implies a procedure to decide whether any given sequence of words (morphemes) is or is not grammatical in L, which implies that L must form a decidable set. Given the well-known problems in devising a *formal* decision grammar for any natural language, say English, one would expect that English is not decidable but merely canonical. Having said this, he proceeds (Beth 1963:21):

> [B]asing ourselves on the foregoing we can […] add the assumption that *the set E1 of all grammatical sentences of English is canonical*. Such a canonical set is, in general, too complex to be decidable; but we can handle it constructively and control it rationally because we can think of the word sequences *w* that belong to it as produced by a fittingly chosen algorithm. […] It would seem to me that Chomsky's work is naturally characterised as an investigation of such an algorithm. What Chomsky usually refers to as a 'model' or a 'grammar' comes very close to an algorithm in the sense of Post; I suspect that systems like those proposed by Chomsky can be transcribed into the forms proposed by Post.

If Beth saw this immediately in 1963, why did Tomalin miss it in 2006? If he had not missed it, he might have developed further thoughts on the question of the decidability of English as a natural language. This question is of considerable interest because if English is not decidable, it is not automatically parsable, which would mean that Categorial Grammar and in general much research in computational linguistics are bound to fail and that the parsing (deciding) ability of natural speakers must be at least partially grounded in extralinguistic information. As it is, there is nothing on the question of the decidability (automatic parsability) of natural languages in *LFS*.

---

1. The algorithmic character of a generative grammar was strongly emphasized in Seuren (1969:26–35).

## 5.   *Positivism, behaviorism and methodology*

But apart from the logico-mathematical background of TGG, there are the philosophical, empirical and methodological aspects. These are almost totally neglected in *LFS*, which biases and ultimately flaws Tomalin's picture of the intellectual climate in which TGG arose. First, there is Tomalin's total neglect of positivism as a philosophy guiding the physical and mathematical sciences and making deep inroads into the human sciences.

In its extreme form, positivism is the philosophy that limits reality to matter and knowledge to observables (see Kolakowski 1972 for an informed survey). During the first half of the 20th century, positivist influences were felt throughout the intellectual community, especially in North America. It enhanced the importance of logic and mathematics because these disciplines were taken to provide the most reliable methods for registering and classifying observables on a large scale and formulating statistical or other inductive regularities. In the philosophy of science, instrumentalism gained the ascendancy over realism. Instrumentalism says that there need not be a hidden or unobservable reality and that scientific theories are mere fabrications to 'control' the data. Realism, by contrast, says that there is an underlying, unobservable but causal, reality, which scientific theories should aim to reconstruct.[2]

Behaviorism, likewise entirely outside Tomalin's range of vision, was a direct application of strict positivist principles to the study of the mind (which was consequently written out of the psychology script). It was also the prime ideology behind Bloomfield's and Harris's notions of natural human language. One only has to read the pages 22 to 26 of Bloomfield's *Language* of 1933 to see the extent to which his views were based on his belief in behaviorism. And anyone reading the Introduction to Harris 1951 will immediately recognize his allegiance to strict positivism and behaviorism, and will see how it determined his method of linguistic analysis. If the axiomatic-deductive method of theorizing is present in the works of Bloomfield, Harris and other linguists of the period, it came from positivist philosophy, not from mathematics, though, of course, mathematics had an important stake in that same positivist philosophy. The reflexes of positivism are clearly recognizable in these linguists' writings, but no mathematics or foundational stuff is to be found there.

The view pressed on the reader by Tomalin (pp. 50, 56, 67, 93–96) that Bloomfield and some of his followers were directly influenced by mathematics and its

---

2.  It was in order to demonstrate my opposition to positivism that I gave my 1969 book the motto "Invisible harmony is stronger than visible harmony", taken from the Greek pre-Socratic philosopher Heraclitus, who was the first to introduce the notion of an unobservable reality causing observable phenomena.

axiomatic-deductive methods is at odds with historical reality to such an extent that it must be considered false. It is true that Bloomfield dabbled in mathematics, but, as Tomalin himself makes clear (p. 96), he himself did not feel that this dabbling stood up to academic standards.

Tomalin sets great store by Bloomfield's "Set of Postulates" of 1926, interpreting it as a direct result of Bloomfield's study of those philosopher-mathematicians who had been proposing the 'axiomatic-deductive method'. He even states (p. 95):

> In essence, as the above passage demonstrates, Bloomfieldian linguistics and the formal sciences were both shaped by [Hilbertian] Formalism during the 1930s, and the effects of this influence are already apparent in Bloomfield's work. For instance, to consider one example, it is well-known that Bloomfield repeatedly expressed scepticism concerning the role of meaning in linguistic theory. […] While there is no doubt that linguistics and psychology were both responsible for determining the direction of Bloomfield's thought in many ways, it is certainly possible that some of his ideas concerning the role of meaning in linguistic theory were directly influenced by his knowledge of Formalism (and/or vice versa), which appeared to advocate the manipulation of meaning-less symbols extracted from their semantic context.

This passage is curious in a variety of ways. First, it shows Tomalin's confusion as regards the notion 'formal' (to which I will revert below). Then, if it is true that "there is no doubt that linguistics and psychology were both responsible for determining the direction of Bloomfield's thought in many ways" (p. 95), and if it is only "possible that some of his ideas concerning the role of meaning in linguistic theory were directly influenced by his knowledge of Formalism" (p. 95), then, one wonders, why are the *undoubted* determiners of "Bloomfield's thought" not given any attention at all in *LFS* while the *possible* ones are dealt with extensively and repetitively? After all, the book carries the subtitle "The Origins of Generative Grammar".

In fact, this lack of proportion brings about a serious distortion of historical truth. Bloomfield's knowledge and understanding of the issues dealt with by the great mathematicians of his day was close to zero — despite his implicit admiration. His rejection of linguistic meaning as a valid object of scientific research and his devout adherence to the Unified Science ideology were entirely due to his positivism (which preached full reduction of all sciences to general physics) and his resulting behaviorism. The historical record is absolutely clear on that (see, for example, Seuren 1998: 197–203). His alleged 'axiomatic' method was not axiomatic at all in the technical sense of the term. Instead, Bloomfield, following his admired colleague, the psychologist Albert Paul Weiss (1879–1931), who published his "One Set of Postulates for a Behavioristic Psychology" in 1925, and followed by Bloch's "A Set of Postulates for Phonemic Analysis" of 1948 for phonology, used

the term *axiomatic* merely as a (then trendy) label for a succinct and systematic statement of the main concepts and tenets of his discipline. There is nothing remotely reminiscent of a formal theory either in Bloomfield's 1926 article or in its counterparts written by Weiss and Bloch. It is, therefore, highly misleading to state, as Tomalin does (p. 50), that "linguists (particularly syntacticians) became increasingly interested in the axiomatic-deductive method during the first decades of the twentieth century."

The statement is more appropriate, however, for some of Bloomfield's followers, in particular Hockett and Harris, the main so-called 'hocus-pocus' linguists (who were given this jocular name because of their predilection for manipulating symbols; see Seuren 1998: 211). Yet even these linguists were interested more in the debates that went on in circles of philosophy of science than in what was happening in the foundational study of formal logic and mathematics. Harris, in particular, was more involved with scientific methodology than with the logico-mathematical aspects of his work, as is evident from the title of his 1951 work.[3]

As far as Chomsky himself is concerned, hardly any mention is made in *LFS* of the fact that Chomsky was educated, by Harris and others, as a behaviorist and that his 1955 *The Logical Structure of Linguistic Theory* was strictly behaviorist (although in the introduction written for its publication in 1975 Chomsky is less than forthcoming about this fact). If anything had contributed to Harris's discovery method and Chomsky's sequel to it, which led directly to TGG, it was behaviorism. In conformity with Tomalin's silence about Chomsky's early behaviorism, there is hardly any mention in *LFS* of Chomsky's conversion to mentalism, first manifested in his 1959 review of Skinner's *Verbal Behavior* of 1957. Yet, given Tomalin's interest in "Chomsky's intellectual development" (p. 125), one would expect at least some discussion of this important aspect of linguistic theory. He might then have shown that Chomsky, upon entering the intellectual scene of Cambridge (Mass.) in the early 1950s, had become associated with a group of angry young Harvard psychologists, comprising men like George A. Miller (b. 1920) and Jerome Bruner (b. 1915), who revolted against the equally Harvard professor Burrhus F. Skinner (1904–1990), then the leading behaviorist in America, and who were in fact preparing the ground for the new development of Cognitive Science, now the dominant paradigm in psychology. And it might have become clear that Chomsky's critique of Skinner and behaviorism in general was based to an overwhelming extent on Karl S. Lashley's (1890–1958) historic paper read at the 1948 Hixon Symposium (Lashley 1951; see Gardner 1985: 28–29), where largely the same arguments are

---

3.  And he was well-placed to entertain optimal contacts with the protagonists in this respect, as his wife Bruria Kaufman, herself a prominent theoretical physicist, was Einstein's (last) assistant at nearby Princeton.

presented, though it is referred to by Chomsky only perfunctorily.[4] Tomalin might then also have mentioned that, as a result of this conversion to mentalism and thus to Cognitive Science, questions of learnability and innateness began to play a decisive role in Chomsky's successive modifications and revisions of TGG.

## 6.   *Zellig Harris buried alive*

In general, *LFS* badly underestimates the internal linguistic forces that led to the genesis of TGG. One should not forget that Chomsky was *nourri dans le sérail* of linguistics, or, to use a different metaphor, that he wears the smell of the linguistics stable. Yes, he was something of a maverick there and hard to handle, but that is where his intellectual home was, not with the mathematicians, logicians, philosophers, psychologists, or what not. It was there that he became a behaviorist and it was there that he became acquainted with the notion of generative grammar and then with that of transformational grammar.

Generative grammar was not conceived, in the early 1950s, because linguists were looking to logic or mathematics but because Zellig Harris drew on the consequences of what he wrote in the last chapter of his *Methods in Structural Linguistics* of 1951. There we read (Harris 1951: 369–370):

> This leads ultimately to sets of few elements having complex definitions but as nearly as possible random occurrence in respect to each other, replacing the original sets of many elements having simple definitions but complexly restricted distribution.

---

4.  In his *Language and Mind* of 1968, Chomsky is more generous with regard to Lashley, but fails to mention his less than generous attitude in his 1959 review of Skinner's *Verbal Behavior*. The following curious passage is found in Chomsky (1968: 2–3; italics mine: PAMS):

> Critical voices, even those that commanded considerable prestige, were simply unheard. For example, Karl Lashley gave a brilliant critique of the prevailing framework of ideas in 1948, arguing that underlying language use — and all organized behavior — there must be abstract mechanisms of some sort that are not analyzable in terms of association and that could not have developed by any such simple means. But his arguments and proposals, though sound and perceptive, had absolutely no effect on the development of the field and went by unnoticed even at his own university (Harvard), then the leading center of psycholinguistic research. *Ten years later Lashley's contribution began to be appreciated, but only after his insights had been independently achieved in another context.*

Even if it is true that Chomsky "independently" came to the insights gained a decade earlier by Lashley, which is unlikely, a more generous reference than the perfunctory note in Chomsky's 1959 would have been appropriate.

Translated back into English from Harris's idiosyncratic jargon, this means that the levelwise induction or discovery process from sounds to phonemes to morphemes to constructions (phrases) to the highest level of sentences, yields a system where the higher elements have less restricted distributions: the higher the element, the less restricted its distribution. Sentences are thus least restricted of all as regards their distribution in texts. They are not totally restricted, but have an "as nearly as possible random occurrence in respect to each other", because there are contextual restrictions on sequences of sentences. The step towards generative grammar is then taken a few pages later (Harris 1951: 372–373):

> The work of analysis leads right up to the statements which enable anyone to synthesize of predict utterances in the language. These statements form a deductive system with axiomatically defined initial elements and with theorems concerning the relations among them. The final theorems would indicate the structure of the utterances of the language in terms of the preceding parts of the system.[5]

One should realize that, according to the Preface, this was written in or before 1947.

What one sees here is, first, that it was Harris's inductive labour that led him to the inverse perspective of a generative production system. Then, one sees that this generative production system is conceived of in terms of an axiomatically organized formal theory in the sense current among logicians of the day, but with the notion of theorem re-interpreted as 'well-formed member of the set of sentences of L' for any such theory (or grammar) of a language L. If anywhere it is here that we must locate the origin of generative grammar, though not yet of transformational grammar. It is thus in linguistics proper that generative grammar was born. Harris's acquaintance with the logico-philosophical and mathematical literature of the day, in particular the writings by Post, Carnap, Tarski, Quine, Goodman, merely made for the spark that lit the creative fire. Tomalin only describes the logico-philosophical and mathematical half of the equation, totally neglecting the linguistics half, with positivism and behaviorism fully integrated into both halves.[6]

---

5.  Harris's indiscriminate use of the terms *sentence* and *utterance* should be seen in the context of his positivist behaviourism: sentences are, after all, abstractions made by linguists, whereas utterances form observable data.

6.  Even Chomsky himself is clear about this, as appears from the following remark (1968: 3):

> … such qualms [regarding positivism and the technological advances of the 1940s] did not go far to dispel the feeling that mathematics, technology, and behavioristic linguistics and psychology were converging on a point of view that was very simple, very clear, and fully adequate to provide a basic understanding of what tradition had left shrouded in mystery.

A similar story can be told for the emergence of the notion of transformation, defining transformational grammar as a specific kind of generative grammar (Seuren 1998: 227–242). Here again, the origin lies in Harris's work, in particular the work he did during the years 1948–1955. Bloomfield, in his *Language* of 1933, also had 'transformations', though the term is never used. So as to simplify the morphological description of the languages concerned, Bloomfield posited underlying forms for German and French adjectives (1933: 217, 219), thereby implicitly introducing the notion of transformation in the now current sense of a procedure to map underlying structures onto surface structures. Chomsky *de facto* followed this procedure in his *Morphophonemics of Modern Hebrew* of 1951, no doubt without realizing, at the time, that he was using what we now call transformations. Harris, likewise, was probably unaware of Bloomfield's 'hidden' notion of transformation.

Harris came upon transformations via his conviction that linguistics should not just study isolated sentences but, rather, coherent sequences of sentences, that is, texts. The reason for this conviction (which Harris shared with European linguists like Louis Hjelmslev or Alan H. Gardiner [(1879–1963)] was that the observational data of linguistics consist not so much of isolated sentences, which are already a linguistic idealization, as of texts or discourses as found in everyday life. His ideal, in this respect, was to gain a grip on the issue of textual coherence. Still being a behaviorist, he could hardly approach this question from a semantic point of view. In fact, whatever existed in the way of semantics in his day did not provide the tools for any such enterprise. Instead, he relied on his old (essentially Bloomfieldian) notion of distribution and sought to approximate textual coherence by defining the distribution of individual sentences in texts. To circumvent the problem that sentences tend to approximate statistical uniqueness and thus zero probability, he turned to lexical role analysis, which tells one what the subject and what the object relation is of NPs with respect to any given transitive verb. A typical passage, in this respect, is the following (Harris 1981 [1952]: 110):

> To this end we would use only those statements of the grammar of the language
> which are true for any sentence of a given form. For example, given any English
> sentence of the form *N1 V N2* (e.g. *The boss fired Jim*), we can get a sentence with
> the noun phrases in the reverse order *N2 – N1* (*Jim – the boss*) by changing the
> suffixes around the verb: *Jim was fired by the boss.* The justification for using such
> grammatical information in the analysis of a text is that since it is applicable to any
> *N1 V N2* sentence in English it must also be applicable to any *N1 V N2* sentence in
> the particular text before us, provided only that it is written in English. The desir
> ability of using such information is that in many cases it makes possible further
> applications of the discourse-analysis method. […] [S]uch use of grammatical
> information does not replace work that could be done by the discourse-analysis

method, nor does it alter the independence of that method. It merely transforms certain sentences of the text into grammatically equivalent sentences […] in such a way that the application of the discourse-analysis method becomes more convenient, or that it becomes possible in particular sections of the text where it was not possible to apply it before.

The point of this exercise is (Seuren 1998: 234) that one expects a distributional definition of sentences to be achieved more easily and more economically if all sentences are presented in a canonical form — that is, in the active mood and with all subjects, objects, etc. filled in explicitly. Passive sentences are thus seen as 'transforms' of their active counterparts. If a text contains, for example, the sentences (a) *Jim was fired by the boss*, (b) *The boss fired the secretary*, and (c) *The director wanted to fire the janitor*, then it makes sense to put the NPs *the boss* and *the director* in one category, and *Jim*, *the secretary* and *the janitor* in another, the former being the firers and the latter the ones fired. This cannot be achieved on distributional grounds if the sentences are taken as they are, but if they are reduced to their 'elementary' forms by undoing the effect of the transformations that have acted upon them, the distributional similarities will stand out more clearly, or so Harris thought.

A few years later, in Harris (1957) which goes back to his presidential address to the Linguistic Society of America in December 1955, the discourse dimension of transformations has been eliminated. We now have a, probably finite, set of underlying 'elementary' or 'kernel' sentences generated by means of a simple phrase-structure generative system. These 'kernel' sentences are input to a transformational component of the grammar, which turns them into surface structures. Transformations are thus no longer 'horizontal' or symmetrical relations between sentences, but asymmetrical, top-down or 'vertical' relations.

In fact, in or around 1957, Harris's and Chomsky's notions of transformation are roughly identical, whereby both *de facto* required that predicate-argument relations remained constant through transformations.[7] Chomsky has always tried to erase his indebtedness to Harris in this respect by either misrepresenting or failing to refer to what Harris had written. The truth is, however, that by 1955 Harris, with Chomsky following in his footsteps, had seen the notion of transformation in the mathematical-logic literature, as is amply demonstrated by Tomalin in *LFS*. But instead of treating a transformation as a logical entailment relation, as in mathematical logic, they transformed transformations into procedures ensuring syntactic well-formedness. This was a totally different interpretation of the general

---

7. Chomsky's statement (1964: 62) that Harris's 1957 notion of transformation is symmetrical is thus clearly false.

and abstract underlying concept of algorithm, inspired not by logic but by Harris's pre-existing notion of transformation that had been developed independently.

Harwood 1955, repeatedly extolled in *LFS*, failed to see this. In the opening passage of his article (Harwood 1955:409) quoted in *LFS* on page 163, Harwood says:

> This paper discusses methods for presenting syntactic information in the form of a calculus, and for measuring its goodness of fit to a language. Given a morpheme list for a language, the aim of a syntactic system is to tell us how to put together the sequences of morphemes which are used as sentences in the language. Such directions we shall call the FORMATION RULES. Additionally, works on syntax usually give a certain amount of information about the equivalences between some sequences and others, e.g. that *John discovered the path = The path was discovered by John*. We shall call such statements TRANSFORMATION RULES. Only a small part of the transformation rules is covered in syntax; some others are discussed in mathematical logic, e.g. Russell's theory of descriptions, procedures of generalization and abstraction. So far in mathematical logic, most attention has been directed to developing formation and transformation rules of artificial languages, and no complete treatment has yet been made of the transformation rules of a natural language. The division between formation and transformation rules can to some extent be altered by altering the units of the system.

This passage shows that Harwood mixed up Harris's early symmetrical 'horizontal' transformations with those of mathematical logic, probably because he took Harris's criterion of semantic equivalence to be a variety of entailment. This in itself is interesting, because when it was proposed in Generative Semantics, after 1964, that syntactic transformations should be semantically invariant, Chomsky, after first endorsing this condition in his *Aspects* of 1965, then violently turned against it, reverting to his original position (and telling the world that he had never believed otherwise).

In any case, it is perfectly clear that Chomsky borrowed his notion of transformation from Harris, even though, in or about 1955, he began going his own way in this respect. Tomalin claims (p. 168) that "the transformations developed by Harris and Chomsky were related to the transformation rules presented in (the English translation of) Carnap *LST* [= *The Logical Syntax of Language* of 1937]." Yet, apart from the question of what is implied by the vague predicate 'related', the focus on Carnap is far too exclusive and is at the expense of the much wider context, including that of linguistics proper, to which these transformations are 'related'.

In this context, it is somewhat curious that Tomalin pays no attention at all to these details or to the question of priority as regards the notion of transformation. Instead, he dismisses the question in the following words (*LFS*, p. 159):

> A considerable amount of attention has been devoted to this topic [i.e. Chomsky's use of the term *transformation*, PAMS] over the years, and the main emphasis has usually been upon establishing the priority either of Harris' or Chomsky's use of the term. While such investigations are undeniably entertaining, a more stimulating assessment can be provided if other sources of influence are considered. Therefore, while Harris' use of the term will be mentioned below, the issue of priority will not really be addressed.

This is an extraordinary statement in a book professing to be about the origins of TGG. Investigations about the question of who was the first to introduce transformations are dismissed as "entertaining" and the issue "will not really be addressed." Full stop. When the reader has recovered from the shock, he or she may recognize the policy generally followed by Chomskyan generativists of either ignoring or ridiculing any argument that would threaten their position. "Never argue with your opponent unless he is a push-over" has been the steadfast motto in those circles since 1970. One detects the same tendency in Tomalin's treatment of Postal (pp. 198–200), whose arguments are dismissed as "diatribes" (in a note on p. 208) without any discussion. Tomalin should realize that in following this policy he risks being branded as an aspiring member of the Chomsky sect, which is no longer a recommendation these days.

## 7.   *The ambiguous term 'formal'*

It thus appears that *LFS* suffers from severe unbalance in its treatment of the origins of early TGG. There is, however, one further point that needs to be mentioned, Tomalin's failure to disentangle the ambiguity of the term *formal* — a failure he shares with many others. In its modern and academically original sense, the predicate *formal* applies to a *calculus*: a formal calculus is an algorithmically organized system allowing for derivations to be carried out solely on the strength of the symbols used and their arrangement with regard to each other. One often finds the term *inscriptional* in this connection. A formal calculus is, in principle, implementable as a computer program requiring no external input from human intelligence. Rosenbloom, in the days when computers were not yet a common commodity, compared a formal calculus to a "happy moron" in the following famous passage (Rosenbloom 1950: 160):

> Another way of looking at these languages is to consider the productions as instructions to a moron, who can scan a string and recognize it as being of a certain form, for producing theorems starting from the axioms. The happy moron can, by merely following the instructions, generate as many theorems as he pleases, and never feels the need for any intelligence in the process. He might just as well be a robot or a machine.

A different sense of *formal* arose in linguistics, where *formal* came to be used as an adjective relating to the notion of 'form' rather than 'meaning', not in a theory but in the object language — an opposition that has dominated linguistic theorizing from Bloomfield to the present day and that resulted not from mathematical Formalism but from positivism. Chomsky confused these two senses of 'formal' when he wrote (Chomsky 1975 [1955]: 83), also quoted in *LFS* on p. 176:

> In the strict sense of the word, an argument, a characterization, a theory, etc. is 'formal' if it deals with form as opposed to meaning, that is, if it deals solely with the shape and arrangement of symbols.

And so did Harwood (1955: 409), quoted above. Both were misled by Carnap, who was himself unclear about the relation between 'syntactically' defined logical deductions and the syntax of natural languages. Carnap wrote (1937: 2):

> For instance, given an appropriate rule, it can be proved that the word-series "Pirots karulize elatically" is a sentence, provided only that "Pirots" is known to be a substantive (in the plural), "karulize" a verb (in the third person plural), and "elatically" an adverb; all of which, of course, in a well-constructed language — as, for example, in Esperanto — could be gathered from the form of the words alone. The meaning of the words is quite inessential to the purpose and need not be known.

Tomalin simply follows suit. On p. 91, having held forth about the notion 'formal' as entertained by Hilbert and Carnap, he refers to Carnap's example "Pirots karulize elatically", adding the comment that Carnap "thus demonstrat[es] that sentences can be exhaustively analysed solely in terms of their formal syntactic structure even if the meaning of the individual words is not known." A few pages later, in the passage cited earlier, he writes (p. 95):

> In essence, as the above passage demonstrates, Bloomfieldian linguistics and the formal sciences were both shaped by [Hilbertian] Formalism during the 1930s, and the effects of this influence are already apparent in Bloomfield's work. For instance, to consider one example, it is well-known that Bloomfield repeatedly expressed scepticism concerning the role of meaning in linguistic theory.

Equating Formalism with scepticism about meaning is seriously mistaken, if only because it would exclude *a priori* any discipline of formal semantics. A theory that is formal in the logico-mathematical sense is itself *uninterpreted* but can be *interpreted upon* any suitable object of enquiry, whether that be a system of law, a system of natural language syntax, or a system of natural language phonology, or a system of natural language semantics. Formal semantics came into being when some philosopher-logicians, in particular Richard Montague (1930–1971), hit upon the idea of defining truth conditions, seen as constituting sentence meanings

in the vein of Alfred Tarski's (1902–1983) 1941 book, in terms of sets of possible worlds, thereby making it possible to set up a complex system of set-theoretically defined functions. Formal semantics consists in the application of such a system to natural-language meaning. This no doubt brilliant achievement is as 'formal' as the study of natural language syntax. Bloomfield's rejection of meaning as an object of scientific enquiry had nothing whatsoever to do with the formalization trend in logic and mathematics as proposed by Hilbert and others, but was entirely motivated by his positivist behaviorism, which rejected anything nonobservable as a possible object of observation and hence as a possible object of scientific investigation.

## 8.  *Some nitpicking*

Finally, a few minor points, just for the record. The name of the Polish logician (Kazimierz) Twardowski (1866–1938) is consistently mis-spelt as *Tawardowski*, even in the alphabetical index. On p. 22 Tomalin gives the name of the 17th-century Italian mathematician as *Buonaventura Cavalieri* (1598–1647), whereas the good man was called (*Francesco*) *Bonaventura Cavalieri*. The bibliographical reference to Weiss (1925) on p. 219 is garbled. And the German plural of *Gegenstand* ("object") is not *\*Gegenstanden*, as Tomalin writes on the pages 75 and 151, but *Gegenstände*.[8]

## 9.  *Conclusion*

In sum, it is appropriate to say that we must be grateful to Tomalin for having called attention to some little known literature, especially Bloomfield's *scripta minora* after 1933 and Chomsky's pre-1957 articles, and also for having highlighted a particular aspect in the origins of TGG. Yet it is equally appropriate to say that he overshot his purpose by grossly inflating the importance of the logico-mathematical background, at the expense of the much wider context in which TGG came into being — a context that included not only linguistics itself but also positivism, behaviorism, psychology, the philosophy of science and perhaps a few other elements of the prevailing intellectual climate of the 1940s and the 1950s.

---

8.  Although I am a little sensitive when such mistakes are made, because German readers might feel slighted, just as English readers would feel dismayed at seeing a German author presenting a similar solecism in a quotation from English, I must immediately confess that I have myself been guilty of a similar offence, when I consistently wrote *President* for what in proper German spelling is *Präsident*, in my book *Chomsky's Minimalism* of 2004.

# REFERENCES

Beth, Evert W. 1963. "Konstanten van het wiskundige denken [Constants in mathematical thinking]". *Mededelingen der Koninklijke Nederlandse Akademie van Wetenschappen. Afdeling Letterkunde. Nieuwe Reeks*, 26:7.231–256. Amsterdam: North-Holland.

Bloch, Bernard. 1948. "A Set of Postulates for Phonemic Analysis". *Language* 24:1.3–46.

Bloomfield, Leonard 1926. "A Set of Postulates for the Science of Language". *Language* 2:1.153–164.

Bloomfield, Leonard. 1933. *Language.* New York: Henry Holt & Co.

Carnap, Rudolf. 1937. *The Logical Syntax of Language.* London: Routledge & Kegan Paul.

Chomsky, Noam. 1957. *Syntactic Structures.* (= *Janua Linguarum*, 4.) The Hague: Mouton.

Chomsky, Noam. 1959a. "On Certain Formal Properties of Grammars". *Information and Control* 2:2.137–167.

Chomsky, Noam. 1959b. Review of B. F. Skinner, *Verbal Behavior* (New York: Appleton–Century–Crofts, 1957). *Language* 35:1.26–58.

Chomsky, Noam. 1964. *Current Issues in Linguistic Theory.* (= *Janua Linguarum*, 38.) The Hague: Mouton.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax.* Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1968. *Language and Mind.* New York: Harcourt, Brace & World.

Gardner, Howard. 1985. *The Mind's New Science: A history of the cognitive revolution.* New York: Basic Books.

Harris, Zellig S. 1951. *Methods in Structural Linguistics.* Chicago: University of Chicago Press.

Harris, Zellig S. 1952. "Discourse Analysis". *Language* 28:1.1–30. (Repr. in Harris 1981:107–142.)

Harris, Zellig S. 1957. "Co-Occurrence and Transformation in Linguistic Structure". *Language* 33:3.283–340. (Repr. in Harris 1981:143–210.)

Harris, Zellig S. 1981. *Papers on Syntax.* Edited by Henry Hiż. Dordrecht: Reidel.

Harwood, F. W[illiam]. 1955. "Axiomatic Syntax: The construction and evaluation of a syntactic calculus". *Language* 31:4.409–413.

Kolakowski, Leszek. 1972. *Positivist Philosophy: From Hume to the Vienna Circle.* Harmondsworth, Middlesex: Penguin Books.

Lashley, Karl S. 1951. "The Problem of Serial Order in Behavior". *Cerebral Mechanisms in Behavior: The Hixon Symposium* ed. by Lloyd A. Jeffress, 112–146. New York: John Wiley & Sons; London: Chapman & Hall.

Post, Emil L. 1944. "Recursively Enumerable Sets of Positive Integers and Their Decision Problems". *Bulletin of the American Mathematical Society* 50.284–316.

Rosenbloom, Paul C. 1950. *The Elements of Mathematical Logic.* New York: Dover.

Seuren, Pieter A. M. 1969. *Operators and Nucleus: A contribution to the theory of grammar.* (= *Cambridge Studies in Linguistics*, 2.) Cambridge: Cambridge University Press.

Seuren, Pieter A. M. 1998. *Western Linguistics: An historical introduction.* Oxford: Blackwell.

Seuren, Pieter A. M. 2004. *Chomsky's Minimalism.* New York: Oxford University Press.

Tarski, Alfred. 1941. *Introduction to Logic and to the Methodology of Deductive Sciences.* New York: Oxford University Press.

Weiss, Albert P. 1925. "One Set of Postulates for a Behavioristic Psychology". *Psychological Review* 32.83–87.

*Author's address:*

Pieter A. M. Seuren
Max Planck Institute for Psycholinguistics
P.O. Box 310
NL-6500 AH Nijmegen
The Netherlands

e-mail: pieter.seuren@mpi.nl